

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:25:34 ; Search time 257.941 Seconds  
(without alignments)  
4108.047 Million cell updates/sec

Title: US-10-617-978-14\_COPY\_62\_240  
Perfect score: 179  
Sequence: 1 cgctgacgtcccggaac.....atgagaacgtgaaggtctga 179

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42.2	23.6	270	5	AAA89399 Scorpion
2	40.8	22.8	270	5	AAA89400 Scorpion
3	40.8	22.8	270	5	AAA89398 Scorpion
4	38	21.2	270	5	AAA89397 Scorpion
C 5	32.6	18.2	336	2	AAT90799 Rat Perse
C 6	32.6	18.2	336	2	AAX60455 Partial s
C 7	32.6	18.2	336	2	AAX60459 WO9914235
C 8	32.6	18.2	391	2	AAX60460 WO9914235
C 9	32.6	18.2	515	8	ABZ53533 Aspergill
C 10	32	17.9	780	4	ABA89006 Escherich
11	32	17.9	2489	6	ABE78877 E. coli C
12	32	17.9	2489	10	ADH80444 Escherich
13	32	17.9	2498	4	ABA89004 Escherich
C 14	32	17.9	2811	4	ABL25174 Drosophil
C 15	31	17.3	4590	5	AAH24065 Yeast AOD
C 16	30.6	17.1	582	6	ABN63438 Human can
17	30.4	17.0	13563	4	ABL06306 Drosophil
18	30.4	17.0	13629	4	ABL06290 Drosophil
19	30.2	16.9	473	3	AAC98441 Human col
20	30.2	16.9	1037	4	AAH34174 Human col

21	30.2	16.9	3144	13	ADR08046
22	30.2	16.9	3378	6	AAK99410
23	30.2	16.9	3379	10	ADB99514
24	30.2	16.9	3379	10	ADB99514
25	30.2	16.9	3379	12	ADO19247
26	30.2	16.9	3379	13	ADP54323
27	30.2	16.9	3379	13	ADP23357
28	30.2	16.9	3402	13	ACN37984
29	30.2	16.9	3406	6	ABL66479
30	30.2	16.9	3406	6	ABK64397
31	30.2	16.9	3406	6	ABN95130
32	30.2	16.9	3445	9	ACH03947
33	30.2	16.9	3445	10	ADJ56503
34	30.2	16.9	3445	12	ADL12594
35	30.2	16.9	3458	6	AAK95018
36	30.2	16.9	3466	8	ACC51032
37	30.2	16.9	3466	8	ABX76294
C 38	30.2	16.9	38643	9	ADA03011
C 39	30.2	16.9	38643	10	ADB72749
C 40	30.2	16.9	38643	10	ADC85491
C 41	30.2	16.9	38643	12	ADM74606
C 42	29.8	16.6	726	6	ABK39306
C 43	29.8	16.6	726	8	ACA11635
C 44	29.8	16.6	726	8	ACA02821
C 45	29.8	16.6	726	10	ADH46863

ALIGNMENTS

RESULT 1  
AAA89399  
ID AAA89399 standard; cDNA; 270 BP.

XX AC AAA89399;  
XX DT 11-SEP-2003 (revised)  
XX DT 23-APR-2001 (first entry)  
XX DE Scorpion sodium channel agonist cDNA clone ibj1c.pk008.f14.  
XX KW Scorpion; venom; toxin; sodium channel agonist; anticonvulsant;  
XX KW nontropic; cerebroprotective; insecticide; ss.  
XX OS Hottentotta judaica.  
XX FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT mat\_peptide 64..267  
FT FT /\*tag= a  
XX PN WO200078957-A2.  
XX PD 28-DEC-2000.  
XX PF 21-JUN-2000; 2000WO-US017048.  
XX PR 22-JUN-1999; 99US-0140410P.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Herrmann R, Lee J, Wong JF;  
XX DR WPI; 2001-050111/06.  
XX DR P-PSDB; AAB20077.

New isolated polynucleotide encoding a scorpion toxin for treating epilepsy, degenerative disorders such as Huntington's disease, and neuronal death following stroke, and for creating plants that are insect-tolerant.  
Claim 1(a); Page 57; 60pp; English.

XX CC The present sequence is that of a portion of the cDNA insert in clone  
CC ibjlc.pk008.f14 that encodes a protein showing 29.6% identity to an  
CC insecticidal toxin of *Orthochirus scrobiculosus*. The clone was isolated  
CC from a scorpion (*Buthotus judaicus*) telson cDNA library. The invention  
CC provides isolated nucleic acid sequences (see AA89386-400) encoding  
CC scorpion toxins (see AAB20064-78) that are sodium channel modifiers. The  
CC invention also relates to the construction of a chimeric gene encoding  
CC all or part of the sodium channel modifier, in sense or antisense  
CC orientation, where expression of the chimeric gene results in production  
CC of altered levels of the sodium channel modifier in a transformed host  
CC cell. Sodium channel modifiers can be used to treat neurological problems  
CC involving abnormal functioning of excitatory amino acid synapses, e.g.  
CC epilepsy, Huntington's disease and neuronal death following stroke.  
CC Genetically engineered recombinant baculoviruses which express protein  
CC toxins capable of incapacitating an insect host can be used as biological  
CC insecticides. The nucleic acids can be used to create transgenic plants  
CC in which sodium channel agonists of the invention are expressed for  
CC improved insect tolerance. (Updated on 11-SEP-2003 to standardise OS  
CC field)

XX SQ Sequence 270 BP; 80 A; 34 C; 60 G; 96 T; 0 U; 0 Other;

Query Match 23.6%; Score 42.2; DB 5; Length 270;  
Best Local Similarity 61.3%; Pred. No. 0.00012;  
Matches 68; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 67 GAGATATCCGACTGCATTAAAGATCTGTCAGAAACCGTGTGGATTACGGGTATTGCT 126  
|||  
Db 128 GTATCATGATTTATTGTCGGACATTTGTAAGTACATGAGTGAATTTGGGTATTGTT 187  
|||

QY 127 ACGCTTCCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTCGAAGGTCT 177  
|||  
Db 188 GGGTCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238  
|||

RESULT 2  
AAA89400  
ID AAA89400 standard; cDNA; 270 BP.  
XX AC AAA89400;  
XX DT 11-SEP-2003 (revised)  
XX DT 23-APR-2001 (first entry)  
XX DE Scorpion sodium channel agonist cDNA clone ibjlc.pk008.119.  
XX KW Scorpion; venom; toxin; sodium channel agonist; anticonvulsant;  
XX KW nootropic; cerebrotective; insecticide; ss.  
XX OS Hottentotta judaica.  
XX PH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 64..267  
FT /\*tag= a  
XX PN WO200078957-A2.  
XX XX  
XX PD 28-DEC-2000.  
XX PF 21-JUN-2000; 2000WO-US017048.  
XX PR 22-JUN-1999; 99US-0140410P.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Herrmann R, Lee J, Wong JF;  
XX XX  
XX DR WPI; 2001-050111/06.  
XX DR P-PSDB; AAB20078.  
XX XX

PT New isolated polynucleotide encoding a scorpion toxin for treating  
PT epilepsy, degenerative disorders such as Huntington's disease, and  
PT neuronal death following stroke, and for creating plants that are insect-  
XX tolerant.  
XX PS Claim 1(a); Page 58; 60pp; English.  
XX CC The present sequence is that of a portion of the cDNA insert in clone  
CC ibjlc.pk008.119 that encodes a protein showing 29.6% identity to an  
CC insecticidal toxin of *Orthochirus scrobiculosus*. The clone was isolated  
CC from a scorpion (*Buthotus judaicus*) telson cDNA library. The invention  
CC provides isolated nucleic acid sequences (see AA89386-400) encoding  
CC scorpion toxins (see AAB20064-78) that are sodium channel modifiers. The  
CC invention also relates to the construction of a chimeric gene encoding  
CC all or part of the sodium channel modifier, in sense or antisense  
CC orientation, where expression of the chimeric gene results in production  
CC of altered levels of the sodium channel modifier in a transformed host  
CC cell. Sodium channel modifiers can be used to treat neurological problems  
CC involving abnormal functioning of excitatory amino acid synapses, e.g.  
CC epilepsy, Huntington's disease and neuronal death following stroke.  
CC Genetically engineered recombinant baculoviruses which express protein  
CC toxins capable of incapacitating an insect host can be used as biological  
CC insecticides. The nucleic acids can be used to create transgenic plants  
CC in which sodium channel agonists of the invention are expressed for  
CC improved insect tolerance. (Updated on 11-SEP-2003 to standardise OS  
CC field)

XX SQ Sequence 270 BP; 82 A; 33 C; 58 G; 97 T; 0 U; 0 Other;

Query Match 22.8%; Score 40.8; DB 5; Length 270;  
Best Local Similarity 61.1%; Pred. No. 0.00038;  
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 70 ATAATCCGACTGCATTAAAGATCTGTCAGAAACCGTGTGGATTACGGGTATTGCTACG 129  
|||  
Db 131 ATCATGATTTATTGTCGGACATTTGTAAGTACATGAGTGAATTTGGGTATTGTTGGG 190  
|||

QY 130 CTTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTCGAAGGTCT 177  
|||  
Db 191 TCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238  
|||

RESULT 3  
AAA89398  
ID AAA89398 standard; cDNA; 270 BP.  
XX AC AAA89398;  
XX XX  
XX DT 11-SEP-2003 (revised)  
XX DT 23-APR-2001 (first entry)  
XX DE Scorpion sodium channel agonist cDNA clone ibjlc.pk006.p4.  
XX KW Scorpion; venom; toxin; sodium channel agonist; anticonvulsant;  
XX KW nootropic; cerebrotective; insecticide; ss.  
XX OS Hottentotta judaica.  
XX PH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 64..267  
FT /\*tag= a  
XX PN WO200078957-A2.  
XX XX  
XX PD 28-DEC-2000.  
XX PF 21-JUN-2000; 2000WO-US017048.  
XX XX  
XX PR 22-JUN-1999; 99US-0140410P.  
XX XX  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.



XX WO9733911-A1.  
 PN 18-SEP-1997.  
 PD 14-MAR-1997; 97WO-US0003461.  
 XX 14-MAR-1996; 96US-00615944.  
 PF (UNIW ) UNIV WASHINGTON.  
 PR Johnson EM, Milbrandt JD, Kotzbauer PT, Lampe PA;  
 XX WPI; 1997-470818/43.  
 XX GDNF-neurturin family related growth factor, Persephin - used to prevent  
 PT or treat cellular, neuronal or non-neuronal, degeneration or  
 PT insufficiency.  
 PS Example 12; Page 157; 228pp; English.  
 XX This sequence comprises a fragment of rat persephin cDNA. Persephin is a  
 CC novel member of the glial-derived neurotrophic factor-neurturin family.  
 CC The clone was obtained by PCR amplification using primers (see AAT90797-  
 CC 98) based on an isolated fragment (see AAT90796) of rat persephin cDNA. A  
 CC full-length rat persephin gene (see AAT90805) was subsequently obtained.  
 CC Mouse and rat persephin polypeptides (see AAM10064-68) and DNA sequences  
 CC encoding them, are used in claimed methods for preventing or treating  
 CC neuronal degeneration, haematopoietic cell degeneration and cardiac  
 CC muscle degeneration or insufficiency  
 XX  
 SQ Sequence 336 BP; 77 A; 91 C; 98 G; 70 T; 0 U; 0 Other;  
 Query Match 18.2%; Score 32.6; DB 2; Length 336;  
 Best Local Similarity 52.6%; Pred. No. 0.37; Mismatches 0; Gaps 0;  
 Matches 71; Conservative 0; Indels 64; Indels 0; Gaps 0;  
 QY 25 CACTTGATCTTCGACAAATACCTACTGTGCGGCCCTTTGGGAGATAATCCGACTGCA 84  
 DB 285 CACTGAGCCTTCAAGGCTGCTCTTGTGTGATCTGAGAGACAGGCTGGCGCC 226  
 QY 85 TTAAGATCTCTCAGAAACACGGTGTGATACGGGTATTCTACGCTTCCAAATGCTGGT 144  
 DB 225 TTCAGCCACCACAGCCCAAGCTGCGGCTGAGAGCTGAGGAGCTGTGCTCAATGTTGGT 166  
 QY 145 GTGAATTTCTGAAGG 159  
 DB 165 GGTCATCAAGGAAGG 151  
 RESULT 6  
 AAX60455/C  
 ID AAX60455 standard; DNA; 336 BP.  
 XX AAX60455;  
 AC  
 XX  
 DT 17-AUG-1999 (first entry)  
 XX Partial sequence of rat persephin cDNA.  
 DE  
 XX Growth factor; GF; persephin; neuron growth; cellular degeneration;  
 XX peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; trauma;  
 KW brain injury; spinal cord injury; nervous system tumour; infection;  
 KW multiple sclerosis; cardiac muscle degeneration; injury; neurotoxin;  
 KW metabolic disease; diabetes; renal dysfunction; neurturin; ss.  
 XX Rattus sp.  
 OS  
 XX  
 XX WO9914235-A1.  
 PN 25-MAR-1999.  
 PD

PF 15-SEP-1998; 98WO-US019163.  
 XX 16-SEP-1997; 97US-00931858.  
 PR (UNIW ) UNIV WASHINGTON.  
 PA Johnson EM, Milbrandt JD, Kotzbauer PT, Lampe PA, Klein R;  
 XX Desauvage F;  
 PI WPI; 1999-244023/20.  
 DR  
 XX New isolated persephin growth factor nucleic acids used to, e.g. promote  
 PT neuronal growth.  
 PT  
 XX Example 12; Page 156; 222pp; English.  
 PS  
 XX The invention relates to a novel isolated and purified growth factor (GF)  
 CC that comprises persephin or a fragment or a conservatively substituted  
 CC variant. The persephin GF polypeptides can promote the survival and  
 CC growth of neurons and non-neuronal cells. The persephin GF polypeptides  
 CC or polynucleotides can be used for preventing or treating cellular  
 CC degeneration or insufficiency, e.g. neuronal degeneration resulting from  
 CC peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease, ischemic stroke,  
 CC acute brain injury, acute spinal cord injury, nervous system tumours,  
 CC multiple sclerosis, or infection, hematopoietic cell degeneration or  
 CC insufficiency resulting from eosinopenia, anemias, thrombocytopenia, or  
 CC stem-cell insufficiencies, cardiac muscle degeneration or insufficiency  
 CC resulting from cardiomyopathy or congestive heart failure. They can also  
 CC be used for treating e.g. peripheral nerve trauma or injury, exposure to  
 CC neurotoxins, metabolic diseases such as diabetes or renal dysfunctions  
 CC and damage caused by infectious agents. The GF can also be used for  
 CC promoting the growth and/or differentiation of a cell in a culture  
 CC medium. The antisense polynucleotides can be used for treating a disease  
 CC condition mediated by expression of persephin by a population of cells.  
 CC The products can also be used for detection and diagnosis  
 XX  
 SQ Sequence 336 BP; 77 A; 91 C; 98 G; 70 T; 0 U; 0 Other;  
 Query Match 18.2%; Score 32.6; DB 2; Length 336;  
 Best Local Similarity 52.6%; Pred. No. 0.37; Mismatches 0; Gaps 0;  
 Matches 71; Conservative 0; Indels 64; Indels 0; Gaps 0;  
 QY 25 CACTTGATCTTCGACAAATACCTACTGTGCGGCCCTTTGGGAGATAATCCGACTGCA 84  
 DB 285 CACTGAGCCTTCAAGGCTGCTCTTGTGTGATCTGAGAGACAGGCTGGCGCC 226  
 QY 85 TTAAGATCTCTCAGAAACACGGTGTGATACGGGTATTCTACGCTTCCAAATGCTGGT 144  
 DB 225 TTCAGCCACCACAGCCCAAGCTGCGGCTGAGAGCTGAGGAGCTGTGCTCAATGTTGGT 166  
 QY 145 GTGAATTTCTGAAGG 159  
 DB 165 GGTCATCAAGGAAGG 151  
 RESULT 7  
 AAX60459/C  
 ID AAX60459 standard; DNA; 336 BP.  
 XX AAX60459;  
 AC  
 XX 17-AUG-1999 (first entry)  
 DT  
 XX WO9914235 Seq ID No: 106.  
 DE  
 XX Growth factor; GF; persephin; neuron growth; cellular degeneration;  
 KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; trauma;  
 KW brain injury; spinal cord injury; nervous system tumour; infection;  
 KW multiple sclerosis; cardiac muscle degeneration; injury; neurotoxin;  
 KW metabolic disease; diabetes; renal dysfunction; neurturin; ss.  
 XX



OS Rattus sp.  
XX WO9914235-A1.  
FN 25-MAR-1999.  
XX 15-SEP-1998; 98WO-US019163.  
XX 16-SEP-1997; 97US-00931858.  
XX (UNIW ) UNIV WASHINGTON.  
XX Johnson EM, Milbrandt JD, Kotzbauer PT, Lampe PA, Klein R;  
PI Desauvage F;  
XX WPI; 1999-244023/20.  
XX New isolated persephin growth factor nucleic acids used to, e.g. promote  
PT neuronal growth.  
XX Example; Page 159; 222pp; English.  
XX The invention relates to a novel isolated and purified growth factor (GF)  
CC that comprises persephin or a fragment or a conservatively substituted  
CC variant. The persephin GF polypeptides can promote the survival and  
CC growth of neurons and non-neuronal cells. The persephin GF polypeptides  
CC or polynucleotides can be used for preventing or treating cellular  
CC degeneration or insufficiency, e.g. neuronal degeneration resulting from  
CC peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's  
CC disease, Parkinson's disease, Huntington's disease, ischemic stroke,  
CC acute brain injury, acute spinal cord injury, nervous system tumours,  
CC multiple sclerosis, or infection, hematopoietic cell degeneration or  
CC insufficiency resulting from eosinopenia, anemias, thrombocytopenia, or  
CC stem-cell insufficiencies, cardiac muscle degeneration or insufficiency  
CC resulting from cardiomyopathy or congestive heart failure. They can also  
CC be used for treating e.g. peripheral nerve trauma or injury, exposure to  
CC neurotoxins, metabolic diseases such as diabetes or renal dysfunctions  
CC and damage caused by infectious agents. The GF can also be used for  
CC promoting the growth and/or differentiation of a cell in a culture  
CC medium. The antisense polynucleotides can be used for treating a disease  
CC condition mediated by expression of persephin by a population of cells.  
CC The products can also be used for detection and diagnosis  
XX Sequence 336 BP; 77 A; 91 C; 98 G; 70 T; 0 U; 0 Other;  
SQ  
Query Match 18.2%; Score 32.6; DB 2; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.37;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
Qy 25 CACTTGATTTCTCGAACAATACCTACTGTGCGGCCCTTTGGGAGATAATCCGACTGCA 84  
Db 285 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGTTGATTCTGAGAGACGAGCTGGCGGCC 226  
Qy 85 TTRAGATCTGTGAGAACACGGTGTGATTACGGGTATTGCTAGCCCTTCCAATGCTGGT 144  
Db 225 TTACGCCACCACACGCCAAGCTCGGCTGAGAGCTGAGGAGCTGCTGCTGCAATGGTGGT 166  
Qy 145 GTGAATTTCTGAAGG 159  
Db 165 GGTCAATCAAGGAGG 151  
RESULT 8  
AAX60460/c  
ID AAX60460 standard; DNA; 391 BP.  
XX AAX60460;  
AC AAX60460;  
XX 17-AUG-1999 (first entry)  
DT  
XX WO9914235 Seq ID No: 107.  
DE  
XX Growth factor; GF; persephin; neuron growth; cellular degeneration;  
KW

KW peripheral neuropathy; amyotrophic lateral sclerosis; ischemic stroke;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; trauma;  
KW brain injury; spinal cord injury; nervous system tumour; infection;  
KW multiple sclerosis; cardiac muscle degeneration; injury; neurotoxin;  
KW metabolic disease; diabetes; renal dysfunction; neurturin; ss.  
XX Unidentified.  
XX WO9914235-A1.  
FN 25-MAR-1999.  
XX 15-SEP-1998; 98WO-US019163.  
XX 16-SEP-1997; 97US-00931858.  
XX (UNIW ) UNIV WASHINGTON.  
XX Johnson EM, Milbrandt JD, Kotzbauer PT, Lampe PA, Klein R;  
PI Desauvage F;  
XX WPI; 1999-244023/20.  
XX New isolated persephin growth factor nucleic acids used to, e.g. promote  
PT neuronal growth.  
XX Example; Page 159; 222pp; English.  
XX The invention relates to a novel isolated and purified growth factor (GF)  
CC that comprises persephin or a fragment or a conservatively substituted  
CC variant. The persephin GF polypeptides can promote the survival and  
CC growth of neurons and non-neuronal cells. The persephin GF polypeptides  
CC or polynucleotides can be used for preventing or treating cellular  
CC degeneration or insufficiency, e.g. neuronal degeneration resulting from  
CC peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's  
CC disease, Parkinson's disease, Huntington's disease, ischemic stroke,  
CC acute brain injury, acute spinal cord injury, nervous system tumours,  
CC multiple sclerosis, or infection, hematopoietic cell degeneration or  
CC insufficiency resulting from eosinopenia, anemias, thrombocytopenia, or  
CC stem-cell insufficiencies, cardiac muscle degeneration or insufficiency  
CC resulting from cardiomyopathy or congestive heart failure. They can also  
CC be used for treating e.g. peripheral nerve trauma or injury, exposure to  
CC neurotoxins, metabolic diseases such as diabetes or renal dysfunctions  
CC and damage caused by infectious agents. The GF can also be used for  
CC promoting the growth and/or differentiation of a cell in a culture  
CC medium. The antisense polynucleotides can be used for treating a disease  
CC condition mediated by expression of persephin by a population of cells.  
CC The products can also be used for detection and diagnosis  
XX Sequence 391 BP; 84 A; 106 C; 117 G; 84 T; 0 U; 0 Other;  
SQ  
Query Match 18.2%; Score 32.6; DB 2; Length 391;  
Best Local Similarity 52.6%; Pred. No. 0.39;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
Qy 25 CACTTGATTTCTCGAACAATACCTACTGTGCGGCCCTTTGGGAGATAATCCGACTGCA 84  
Db 337 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGTTGATTCTGAGAGACGAGCTGGCGGCC 278  
Qy 85 TTRAGATCTGTGAGAACACGGTGTGATTACGGGTATTGCTAGCCCTTCCAATGCTGGT 144  
Db 277 TTACGCCACCACACGCCAAGCTCGGCTGAGAGCTGAGGAGCTGCTGCTGCAATGGTGGT 218  
Qy 145 GTGAATTTCTGAAGG 159  
Db 217 GGTCAATCAAGGAGG 203  
RESULT 9  
ABZ53533  
ID ABZ53533 standard; cDNA; 515 BP.  
XX ABZ53533;  
AC ABZ53533;



XX WO200259320-A2.  
XX 01-AUG-2002.  
XX 19-OCT-2001; 2001WO-US046833.  
XX 19-OCT-2000; 2000US-0242412P.  
XX (WISC.) WISCONSIN ALUMNI RES FOUND.  
XX Blattner FR, Welch RA, Burland VD;  
XX WPI; 2002-691532/74.  
XX New DNA sequences of the pathogenic *Escherichia coli* CFT073 strain,  
XX useful for preventing or treating *E. coli* CFT073 infection in humans or  
XX livestock.  
XX Claim 1; Page 141-142; 765pp; English.  
XX The present invention relates to polynucleotide sequences from the genome  
XX of the pathogenic *Escherichia coli* strain CFT073. Almost all the  
XX sequences present in *E. coli* CFT073 are absent in the previously  
XX sequenced laboratory strain K-12. The polynucleotide sequences of the  
XX CFT073 infection in humans or livestock. The polynucleotide sequences are  
XX useful for preventing urinary tract infections and pyelonephritis.  
XX Likewise, the polypeptides encoded by the different open reading frames  
XX (ORF1-5) are useful for generating a vaccine against uropathogenic *E. coli*  
XX strains. ABS78834-ABS79085 represent genomic sequences from *E. coli*  
XX strain CFT073  
XX  
XX Sequence 2489 BP; 668 A; 563 C; 578 G; 677 T; 0 U; 3 Other;  
Query Match 17.9%; Score 32; DB 6; Length 2489;  
Best Local Similarity 58.3%; Pred. No. 1.3;  
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
Qy 33 TCTTCCGACATACCTACCTGTGCGCCCTTTGGAGATATCCGGACTGCATTAAAGATC 92  
Db 901 TCTGCGCCCAATACGTCCCTGACCATCAGAGCCGAGTGTAATAATGATGGCATCAATATC 960  
Qy 93 TGTGAGAAACACGGTGTGGATTACGGGTATTGCTTAC 128  
Db 961 TTTCGGTAAATCGGTAAAGTGTTCAGGTATTGAAAC 996  
RESULT 12  
ADH80444  
ID ADH80444 standard; DNA; 2489 BP.  
XX AC ADH80444;  
XX 22-APR-2004 (first entry)  
XX *Escherichia coli* CFT073 genome contig #44.  
XX ds; gene; *Escherichia coli*; CFT073; *Escherichia coli* CFT073 infection.  
XX *Escherichia coli*; CFT073.  
XX US2003165870-A1.  
XX 04-SEP-2003.  
XX 01-MAR-2002; 2002US-00085959.  
XX 01-MAR-2002; 2002US-00085959.  
XX (BLAT/) BLATTNER F R.  
XX (WELC/) WELCH R A.  
XX (BURL/) BURLAND V D.

XX Blattner FR, Welch RA, Burland VD;  
XX WPI; 2003-863698/80.  
XX New nucleic acid of *Escherichia coli* CFT073, useful for preparing a  
XX composition for diagnosing, treating or preventing infection caused by  
XX *Escherichia coli* CFT073.  
XX Claim 1; SEQ ID NO 44; 4pp; English.  
XX The invention relates to an isolated *Escherichia coli* CFT073 nucleic acid  
XX molecule. The nucleic acid is useful for preparing a composition for  
XX diagnosing, treating or preventing infection caused by *Escherichia coli*  
XX CFT073. The present sequence represents a contig of the *Escherichia coli*  
XX CFT073 genome that is not present in *Escherichia coli* K-12.  
XX Sequence 2489 BP; 668 A; 563 C; 578 G; 677 T; 0 U; 3 Other;  
Query Match 17.9%; Score 32; DB 10; Length 2489;  
Best Local Similarity 58.3%; Pred. No. 1.3;  
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
Qy 33 TCTTCCGACATACCTACCTGTGCGCCCTTTGGAGATATCCGGACTGCATTAAAGATC 92  
Db 901 TCTGCGCCCAATACGTCCCTGACCATCAGAGCCGAGTGTAATAATGATGGCATCAATATC 960  
Qy 93 TGTGAGAAACACGGTGTGGATTACGGGTATTGCTTAC 128  
Db 961 TTTCGGTAAATCGGTAAAGTGTTCAGGTATTGAAAC 996  
RESULT 13  
ABA89004  
ID ABA89004 standard; DNA; 2498 BP.  
XX AC ABA89004;  
XX 11-FEB-2002 (first entry)  
XX *Escherichia coli* polynucleotide SEQ ID NO 578.  
XX *Escherichia coli*; B2/D-A; antiinflammatory; antibacterial;  
XX immunosuppressive; extra-intestinal infection; phylogeny; meningitis;  
XX systemic infection; non-diarrhoeal infection; septicemia;  
XX pyelonephritis; antibiotic resistance; ds.  
XX *Escherichia coli*.  
XX WO200166572-A2.  
XX 13-SEP-2001.  
XX 12-MAR-2001; 2001WO-EP003445.  
XX 10-MAR-2000; 2000FR-00003145.  
XX 02-FEB-2001; 2001FR-00001449.  
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;  
XX WPI; 2001-550253/61.  
XX A library of DNA fragments of *Escherichia coli* strains for the phylogenetic  
XX determination of a given strain comprises polynucleotides of nature B2/D+  
XX A-.  
XX Example 6; Fig 6; 646pp; English.  
XX The invention relates to a library of DNA fragments of *Escherichia coli*  
XX strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533) and  
XX encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature

B2/D/A-. The polynucleotides have potential antiinflammatory, antibacterial and immunosuppressive activity as part of pharmaceutical compositions used to treat, palliate or prevent extra-intestinal E. coli infections. The polypeptides are useful for determining the phylogenetic group of a given E. coli strain. These polypeptides can detect and treat an undesired development of E. coli, particularly an extra-intestinal infection that include systemic and non-diarrhoeal infections such as septicæmia, pyelonephritis and meningitis this is particularly advantageous as bacterial resistance is increasing with the more frequent use of broad spectrum antibiotics

Query Match 17.9%; Score 32; DB 4; Length 2498;  
Best Local Similarity 58.3%; Pred. No. 1.3;  
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 33 TCITCCGACAAATACCTACCTGTGTGGCCCTTTGGGAGATATCCGGACTGCATTAAATC 92  
Db 906 TCTGCGCCCAATACGTCCTCCCTGACCATCAGAGCCGAGTGAATAATGATGGCATCATATC 965  
Qy 93 TGTGAGAAACACGCTGTGGATTACGGGTATTGCTAC 128  
Db 966 TTTCCGTAATTCGGTAAGTGTTTCAGGTATTGAAAC 1001

RESULT 14  
ABL25174/c  
ID ABL25174 standard; DNA; 2811 BP.  
XX  
AC ABL25174;  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 26995.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
FI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
WPI; 2001-656860/75.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions.  
XX  
PS Claim 1; SEQ ID NO 26995; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC AB570272). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX

Qy	Sequence	2811 BP	719 A	640 C	660 G	792 T	0 U	0 Other									
Query Match	17.9%	Score 32	DB 4	Length 2811													
Best Local Similarity	58.3%	Pred. No. 1.4															
Matches	56	Conservative	0	Mismatches	40	Indels	0	Gaps 0									
Qy	1	CGGCTGAGTCCTCCGGAAACTACCCACTTGATTCCTCCGACAAATACCTACCTGTGCGCCC	60														
Db	2672	CTGTGACGCCCGACATATTACTCCACGATACTTGGGCATACAACTTCATCGCTCTGC	2613														
Qy	61	CTTTGGGAGATATCCGGACTGCTTAAGATCTGTC	96														
Db	2612	CCTGTGAGCTTATCTCGCTCGCTTAAAGTTGGGC	2577														
RESULT 15																	
AAH24065/C																	
ID	AAH24065	standard	DNA	4590	BP												
XX	AAH24065																
XX	29-AUG-2001	(first entry)															
XX	Yeast AOD9604-associated DNA sequence	SEQ ID NO:1.															
XX	Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;																
XX	modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;																
XX	lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;																
XX	functional food; transgenic yeast; fat/lean ratio; food use; ds.																
OS	Saccharomyces cerevisiae.																
XX	Key	Location/Qualifiers															
XX	misc_feature	10															
FT		/tag= a															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3617															
FT		/tag= b															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3649															
FT		/tag= c															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3679															
FT		/tag= d															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3819															
FT		/tag= e															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3862															
FT		/tag= f															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3864															
FT		/tag= g															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3888															
FT		/tag= h															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3890															
FT		/tag= i															
FT		/note= "Represented as * in the specification"															
FT	misc_feature	3912															
FT		/tag= j															
FT		/note= "Represented as * in the specification"															



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 15:22:21 ; Search time 177 Seconds  
(without alignments)  
167.800 Million cell updates/sec

Title: US-10-617-978-20

Perfect score: 343

Sequence: 1 ADVPGNYPLDSSDNTLYCAP.....GYCAFQWCFLKDNVKV 58

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	64.7	58	1 BIRT PART	P58752 parabuthus
2	188	54.8	80	2 Q6WJF5	Q6WJF5 mesobuthus
3	127	37.0	67	1 SCXC_CENNO	P63019 centrurioide
4	123	35.9	64	2 Q6VAY9	Q6VAY9 centrurioide
5	119.5	34.8	65	1 SCXM_CENSC	P56646 centrurioide
6	117	34.1	64	2 Q6VAY8	Q6VAY8 centrurioide
7	117	34.1	87	1 SCX2_CENSC	P01493 centrurioide
8	117	34.1	87	1 SCX3_CENSC	P01494 centrurioide
9	116	33.8	64	2 Q6VAY4	Q6VAY4 centrurioide
10	116	33.8	64	2 Q6VAY7	Q6VAY7 centrurioide
11	115	33.5	64	2 Q68PG2	Q68PG2 centrurioide
12	115	33.5	73	2 Q68PG2	Q68PG2 centrurioide
13	115	33.5	85	1 SCX7_CENLL	P59865 centrurioide
14	113	32.9	66	2 Q68PG4	Q68PG4 centrurioide
15	113	32.9	67	2 Q68PG7	Q68PG7 centrurioide
16	113	32.9	87	1 SCXB_CENLL	Q721K7 centrurioide
17	111	32.4	87	1 SCXB_CENNO	P45663 centrurioide
18	110	32.1	66	1 SCX2_CENSC	P08900 centrurioide
19	110	32.1	87	1 SCX3_CENSC	Q95Wd2 centrurioide
20	109	31.8	64	2 Q6VAY3	Q6VAY3 centrurioide
21	109	31.8	66	1 SCXR_CENLL	P45667 centrurioide
22	109	31.8	87	1 SCXX_CENNO	P45664 centrurioide
23	108	31.5	62	1 SCXB_CENNO	Q9tW10 centrurioide
24	108	31.5	66	1 SCX2_CENLL	P59898 centrurioide
25	108	31.5	84	1 SCX2_CENNO	P01495 centrurioide
26	107	31.2	64	2 Q6VAY1	Q6VAY1 centrurioide
27	107	31.2	64	2 Q6VAY2	Q6VAY2 centrurioide
28	107	31.2	64	2 Q6VAY6	Q6VAY6 centrurioide
29	107	31.2	66	1 SCX1_CENII	P59897 centrurioide
30	107	31.2	66	1 SCX1_CENLI	P18926 centrurioide
31	107	31.2	66	1 SCX1_CENLL	P45666 centrurioide

32	107	31.2	85	1 SCCE_CENLL	Q721K6 centrurioide
33	107	31.2	87	1 SC2B_CENLL	P59899 centrurioide
34	107	31.2	87	1 SCX4_CENNO	P45662 centrurioide
35	107	31.2	87	1 SCXC_CENLL	Q7Y661 centrurioide
36	106.5	31.0	87	1 SCX1_CENSC	P01492 centrurioide
37	106	30.9	64	2 Q6VAY5	Q6VAY5 centrurioide
38	106	30.9	69	2 Q68PG9	Q68PG9 centrurioide
39	106	30.9	69	2 Q68PH2	Q68PH2 centrurioide
40	106	30.9	85	1 SCX5_CENLL	Q721K5 centrurioide
41	106	30.9	87	2 Q68PH1	Q68PH1 centrurioide
42	105	30.6	69	2 Q68PG8	Q68PG8 centrurioide
43	105	30.6	69	2 Q68PH0	Q68PH0 centrurioide
44	105	30.6	87	2 Q68PH3	Q68PH3 centrurioide
45	105	30.6	87	2 Q68PH4	Q68PH4 centrurioide

## ALIGNMENTS

### RESULT 1

ID BIRT PART STANDARD; PRT; 58 AA.  
AC P58752;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Birtoxin.  
OS Parabuthus transvaalicus (South African fattenail scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Butthida; Butthoidea; Butthidae; Parabuthus.  
RX NCBI\_TaxID=170972;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=2151824; PubMed=11606203;  
RA Inceoglu B., Lango J., Wu J., Hawkins P., Southern J., Hammock B.D.;  
RT "Isolation and characterization of a novel type of neurotoxic peptide from the venom of the South African scorpion Parabuthus transvaalicus (Butthidae).";  
RL Eur. J. Biochem. 268:5407-5413(2001).  
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation of the activated channels, thereby blocking neuronal transmission (By similarity). Moderately toxic, but very high abundant. Lethal to mice. Do not target reptilian channels.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -!- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family. HSP; P41964; lMYN.  
DR InterPro; IPR002061; Scorpion\_toxinL.  
DR Pfam; PF00537; Toxin\_3; 1.  
DR ProDom; PD000308; Scorpion\_toxinL; 1.  
KW Direct protein sequencing; Ionic channel inhibitor; Neurotoxin;  
KW Sodium channel inhibitor; Toxin.  
FT DISULFID 18 41 By similarity.  
FT DISULFID 27 46 By similarity.  
FT DISULFID 31 48 By similarity.  
SQ SEQUENCE 58 AA; 6548 MW; F5C9EA421959D096 CRC64;

Query Match 64.7%; Score 222; DB 1; Length 58;

Best Local Similarity 60.3%; Pred. No. 2.4e-18;

Matches 35; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 ADVPGNYPLDSSDNTLYCAPLGDNDPDKIKCKHGVDYCYCYAFQWCFLKDNVKV 58

DB 1 ADVPGNYPLDGDNTYKCFLLGGNECLNVCKLHGQVGYCYAKWCCEYLEDDKDSV 58

### RESULT 2

Q6WJF5 PRELIMINARY; PRT; 80 AA.

ID Q6WJF5

AC Q6WJF5;

DT 05-JUL-2004 (TremblRel. 27, Created)

DT 05-JUL-2004 (TremblRel. 27, Last sequence update)



DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Toxin KBR precursor.  
OS Mesobuthus marseus (Manchurian scorpion) (Buthus marseus).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthida; Buthoidea; Buthidae; Mesobuthus.  
OX NCBI\_TaxID=34649;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Venom;  
RQ JIANG D., CAO Z., LI W.;  
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
RL -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -1- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.  
DR EMBL; A2282464; A222733.1; -.  
DR HSSP; P01493; 1JZA.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR002061; Scorpion\_toxinL.  
DR Pfam; PF00537; Toxin\_3; 1.  
DR ProDom; PD000908; Scorpion\_toxinL; 1.  
DR Signal.  
KW SIGNAL.  
FT CHAIN 1 22 Potential.  
FT CHAIN 23 80 toxin KBT.  
FT CHAIN 80 AA; 9274 MW; 5298A578FA62905 CRC64;  
SQ SEQUENCE 80 AA; 9274 MW; 5298A578FA62905 CRC64;  
  
Query Match 54.8%; Score 188; DB 2; Length 80;  
Best Local Similarity 54.4%; Pred. No. 3e-14;  
Matches 31; Conservative 9; Mismatches 17; Indels 0; Gaps 0;  
  
QY 2 DVEGNYPLDSSDNTYLCAPIGDNPDCIKICQKHGVDYGYCYAFQWCCEFLKDNV 58  
DB 24 DVEGNYPTNAYGNKYCYTILGENEYCRKICKLHGVTGYCYNSRCWCEKLEDKDV 80  
  
RESULT 3  
SCXC\_CENNO STANDARD; PRT; 67 AA.  
ID SCXC\_CENNO  
AC Pe3019;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Neurotoxin Cn12.  
OS Centruroides noxius (Mexican scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthida; Buthoidea; Buthidae; Centruroides.  
OX NCBI\_TaxID=6876;  
RN [1]  
RP SEQUENCE, MASS SPECTROMETRY, AND STRUCTURE BY NMR.  
RC TISSUE=Venom;  
RX PubMed=15182366; DOI=10.1111/j.1432-1033.2004.04181.x;  
RA del Rio-Portilla F., Hernandez-Marin E., Pimental G., Coronas F.V.,  
Zamudio F.Z., Rodriguez de la Vega R.C., Wanke E., Possani L.D.;  
RT "NMR solution structure of Cn12, a novel peptide from the Mexican  
scorpion Centruroides noxius with a typical beta-toxin sequence but  
with alpha-like physiological activity.";  
RL Eur. J. Biochem. 271:2504-2516(2004).  
CC -1- FUNCTION: Binds, in vitro, to sodium channels and inhibits the  
inactivation of the activated channels. Seems not toxic to mice,  
crickets and sweet-water shrimps.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- MASS SPECTROMETRY: MW=7139.5; RANGE=1-67; NOTE=Ref.1.  
CC -1- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family. Beta-  
toxin subfamily.  
PDB; 1PB4; NMR; A1-67.  
KW 3D-structure; Direct protein sequencing; Ionic channel inhibitor;  
KW Neurotoxin; Sodium channel inhibitor; Toxin.  
FT DISULFID 11 65  
FT DISULFID 15 40  
FT DISULFID 25 45  
FT DISULFID 29 47  
SQ SEQUENCE 67 AA; 7148 MW; B7BCA58566FEC0EC CRC64;

Query Match 37.0%; Score 127; DB 1; Length 67;  
Best Local Similarity 46.2%; Pred. No. 3.1e-07;  
Matches 24; Conservative 10; Mismatches 16; Indels 2; Gaps 2;  
  
QY 7 YPLDSSDNTYLCAPIGDNPDCIKICQKHGVDYGYCYAFQWCCEFLKDNV 56  
DB 4 YPLASNGCKFGCSGLGNNPTCNHVCERKAGSDYGYAWTCYCEHVAEGTV 55  
  
RESULT 4  
Q6VAY9 PRELIMINARY; PRT; 64 AA.  
ID Q6VAY9  
AC Q6VAY9;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Beta-toxin (Fragment).  
OS Centruroides sculpturatus (Bark scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthida; Buthoidea; Buthidae; Centruroides.  
OX NCBI\_TaxID=218467;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhu S.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -1- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.  
DR EMBL; AY351301; AAR08036.1; -.  
DR HSSP; P01493; 1JZA.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO; GO:0005515; P:protein binding; IEA.  
DR GO; GO:0006952; P:defense response; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003614; Knt1.  
DR InterPro; IPR001219; Neurotoxin.  
DR InterPro; IPR002061; Scorpion\_toxinL.  
DR Pfam; PF00537; Toxin\_3; 1.  
DR PRINTS; PR00285; SCORPNTOXIN.  
DR PRINTS; PR00284; TOXIN.  
DR ProDom; PD000908; Scorpion\_toxinL; 1.  
DR SMART; SM00505; Knt1; 1.  
DR NON\_TER 1 1  
FT NON\_TER 64 64  
SQ SEQUENCE 64 AA; 6959 MW; 730092293C661C41 CRC64;  
  
Query Match 35.9%; Score 123; DB 2; Length 64;  
Best Local Similarity 46.9%; Pred. No. 8.7e-07;  
Matches 23; Conservative 6; Mismatches 18; Indels 2; Gaps 1;  
  
QY 7 YPLDSSDNTYLCAPIGDNPDCIKICQKHGVDYGYCYAFQWCCEFLKDNV 53  
DB 4 YLVNSTGCKYGLKLGNEGCKEAKNQGSGYGYCYAFQWCCEGLPE 52  
  
RESULT 5  
SCXC\_CENNO STANDARD; PRT; 65 AA.  
ID SCXC\_CENNO  
AC P5646;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Toxin Cse M1 (CseM1) (CseM1).  
OS Centruroides sculpturatus (Bark scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthida; Buthoidea; Buthidae; Centruroides.  
OX NCBI\_TaxID=218467;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=93162874; PubMed=1286943;  
RX Pete M.J., Conlon J.M., Murphy R.F.;

Best Local Similarity 46.9%; Pred. No. 4.4e-06;  
Matches 23; Conservative 5; Mismatches 19; Indels 2; Gaps

Neurotoxin; Signal; Sodium channel inhibitor; Toxin.	
FT SIGNAL	1 19
FT PEPTIDE	20 85
	Neurotoxin 2.

```
FT DISULFID 31 84
FT DISULFID 35 60
FT DISULFID 44 65
FT DISULFID 48 67
FT VARIANT 6 6
FT VARIANT 11 14
FT VARIANT 20 20
FT VARIANT 30 30
FT VARIANT 51 51
FT CONFLICT 44 46
FT CONFLICT 83 84
FT STRAND 22 23
FT STRAND 25 25
FT TURN 27 29
FT STRAND 32 32
FT STRAND 34 34
FT STRAND 39 39
FT HELIX 42 49
FT TURN 51 53
FT STRAND 57 60
FT STRAND 64 68
FT TURN 72 73
FT STRAND 79 79
FT TURN 80 81
FT STRAND 82 82
SQ SEQUENCE 87 AA; 9520 MW; 760810C65269B74E CRC64;

Query Match 34.1%; Score 117; DB 1; Length 87;
Best Local Similarity 52.5%; Pred. No. 6e-06;
Matches 21; Conservative 3; Mismatches 14; Indels 2; Gaps 1;

QY 16 YLCAPLGNDPCIKIC--QKHGVVDYGYCYAPQCWCFELKD 53
DB 33 YGCLGLGENEGDKCKAKNQGSYGYCYAFACWCEGLPE 72

RESULT 8
SCX3_CENSC
ID SCX3_CENSC STANDARD; PRT; 87 AA.
AC P01494; Q95WB9; Q95WCO;
DT 21-JUL-1986 (Rel. 01, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Neurotoxin 3 precursor (CSEv3) (CSE v3).
OS Centruroides sculpturatus (Bark scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthida; Buthoidea; Buthidae; Centruroides.
ON NCBI_TaxID=218467;
RX SEQUENCE FROM N.A. (ISOFORMS CSEV3B AND CSEV3B*).
RC TISSUE=Venom gland;
RA MEDLINE=21486965; PubMed=11600153; DOI=10.1016/S0041-0101(01)00174-X;
RA Corona M., Valdez-Cruz N.A., Merino E., Zurita M., Posasani L.D.;
RT "Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
RT that recognize Na(+)-channels.";
RL Toxicol 39:1893-1898(2001).
RN [2]
RN SEQUENCE OF 20-84 (ISOFORM CSEV3A).
RC TISSUE=Venom;
RX MEDLINE=75163395; PubMed=4450885;
RA Babin D.R., Watt D.D., Goos S.M., Mlejnek R.V.;
RT "Amino acid sequences of neurotoxic protein variants from the venom of
RT Centruroides sculpturatus Ewing.";
RL Arch. Biochem. Biophys. 164:694-706(1974).
RN [3]
RN REVISIONS.
RX MEDLINE=82200153; PubMed=7080025; DOI=10.1016/0041-0101(82)90137-4;
RA Shown A., Mole J.;
RL Unpublished results, cited by:
RL Pontecilla-Camps J.-C., Almassy R.J., Suddath F.L., Bugg C.E.;
RL Toxicol 20:1-7(1982).
RN [4]
RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS), AND DISULFIDE BONDS.

RX MEDLINE=82200153; PubMed=7080025; DOI=10.1016/0041-0101(82)90137-4;
RA Pontecilla-Camps J.-C., Almassy R.J., Suddath F.L., Bugg C.E.;
RT "The three-dimensional structure of scorpion neurotoxins.";
RL Toxicol 20:1-7(1982).
RN [5]
RX X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RL MEDLINE=81101077; PubMed=6256740;
RA Pontecilla-Camps J.-C., Almassy R.J., Suddath F.L., Watt D.D.,
RA Bugg C.E.;
RT "Three-dimensional structure of a protein from scorpion venom: a new
RT structural class of neurotoxins.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:6496-6500(1980).
RN [6]
RX X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
RL MEDLINE=92395664; PubMed=1522588;
RA Zhao B., Carson M., Ealick S.E., Bugg C.E.;
RT "Structure of scorpion toxin variant-3 at 1.2-A resolution.";
RL J. Mol. Biol. 227:239-252(1992).
RN [7]
RX STRUCTURE BY NMR.
RL MEDLINE=94062845; PubMed=8243479;
RA Lee W., Moore C.H., Watt D.D., Krishna N.R.;
RT "Aminoacyl chloromethanes as tools to study the requirements of NADPH
RT oxidase activation in human neutrophils.";
RL Eur. J. Biochem. 218:89-95(1993).
CC -|- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC of the activated channels, thereby blocking neuronal transmission.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -|- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family. Beta-
CC toxin subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF338461; AAL23429.1; -.
CC EMBL; AF338462; AAL23430.1; -.
CC PDB; 2SN3; X-ray; @=1-65.
CC InterPro; IPR001219; Neurotoxin.
CC InterPro; IPR002061; Scorpion_toxinL.
CC Pfam; PF00537; Toxin_3; 1.
CC PRINTS; PR00285; SCORPNTOXIN.
CC PRINTS; PR00284; TOXIN.
CC ProDom; PD000908; Scorpion_toxinL; 1.
CC 3D-structure; Amidation; Direct protein sequencing;
CC Ionic channel inhibitor; Neurotoxin; Signal; Sodium channel inhibitor;
CC Toxin.
CC SIGNAL 1 19 Neurotoxin 3.
CC PEPTIDE 20 84 Cysteine amide (G-85 provides amide
CC MOD_RES 84 84 group) (potential).
CC
CC DISULFID 31 84
CC DISULFID 35 60
CC DISULFID 44 65
CC DISULFID 48 67
CC VARIANT 6 6
CC VARIANT 26 26
CC VARIANT 29 29
CC VARIANT 35 35
CC VARIANT 46 46
CC STRAND 22 23
CC STRAND 25 25
CC TURN 27 29
CC STRAND 32 32
CC STRAND 34 34
CC STRAND 39 39
CC HELIX 42 49
CC TURN 51 53
```

OX	NCBI_TaxID=218467;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Zhu S.;
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC	-1- SUBCELLULAR LOCATION: Secreted (By similarity)
CC	-1- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.
DR	EMBL; AY351306; AR08041.1; -.
DR	HSSP; P01493; 1JZA.
DR	GO; GO:0005576; C:extracellular; IEA.
DR	GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR	GO; GO:0005515; F:protein binding; IEA.
DR	GO; GO:0006952; P:defense response; IEA.
DR	GO; GO:0009405; P:pathogenesis; IEA.
DR	InterPro; IPR003614; Knot1.
DR	InterPro; IPR001219; Neurotoxin.
DR	Pfam; PF00537; Toxin_3; 1.
DR	PRINTS; PR00285; SCORPNTXIN.
DR	PRINTS; PR00284; TOXIN.
DR	PRODOM; PD000908; Scorpion_toxinL; 1.
DR	SMART; SM00505; Knot1; 1.
FT	NON_TER 1 1
FT	NON_TER 64 64
SQ	SEQUENCE 64 AA; 7137 MW; 8BEFED934D78C47 CRC64;
Query Match	33.8%; Score 116; DB 2; Length 64;
Best Local Similarity	46.9%; Pred.No.5.7e-06;
Matches	23; Conservative 4; Mismatches 20; Indels 2; Gaps 1;
Qy	7 YPLDSNDTYVLCAPGDNPDCIKIC--QRHGVDYGYCYAFQCWCEFLKD 53   :   :   :   :   :   :   :   :   :   :   :   :   : 4 YLVKSDGCKYDYCFWLGENEGCDKECAKNQGSGSYGYCYAFACWCEGLPE 52
Dd	
RESULT 11	
Q6v4y7	
ID	Q6V4Y7 PRELIMINARY; PRT; 64 AA.
AC	Q6V4Y7;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Beta-toxin (Fragment).
OS	Centruroides noxius (Mexican scorpion).
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC	Buthida; Buthoidea; Buthidae; Centruroides.
OX	NCBI_TaxID=6878;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Zhu S.;
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC	-1- SUBCELLULAR LOCATION: Secreted (By similarity)
CC	-1- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.
DR	EMBL; AY351303; AR08038.1; -.
DR	HSSP; P01493; 1JZA.
DR	GO; GO:0005576; C:extracellular; IEA.
DR	GO; GO:0008200; F:ion channel inhibitor activity; IEA.
DR	GO; GO:0005515; F:protein binding; IEA.
DR	GO; GO:0006952; P:defense response; IEA.
DR	GO; GO:0009405; P:pathogenesis; IEA.
DR	InterPro; IPR003614; Knot1.
DR	InterPro; IPR001219; Neurotoxin.
DR	Pfam; PF00537; Toxin_3; 1.
DR	PRINTS; PR00285; SCORPNTXIN.
DR	PRINTS; PR00284; TOXIN.
DR	PRODOM; PD000908; Scorpion_toxinL; 1.
DR	SMART; SM00505; Knot1; 1.
FT	NON_TER 1 1
FT	NON_TER 64 64
SQ	SEQUENCE 64 AA; 7050 MW; DEBFAD255D08A8D CRC64;
Query Match	33.5%; Score 115; DB 2; Length 64;

```

Best Local Similarity 46.9%; Pred. No. 7.4e-06;
Matches 23; Conservative 5; Mismatches 19; Indels 2; Gaps 1;

QY 7 YPLDSSDNTYLCAPLGNPDCKIKC--QRKHGVDYGYCYAFQWCCEFLKD 53
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 YLVNSTGCKYGCFLWLGKNEGCDCKEAKNQGGSYGYCYAFGCWCEGLPE 52

RESULT 12
Q68PG2 PRELIMINARY; PRT; 73 AA.
AC Q68PG2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Cex13 neurotoxin (Fragment).
GN Name=Cex13;
OS Centruroides exilicauda (Bark scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthida; Buthoidea; Buthidae; Centruroides.
OX NCBI_TaxID=6879;
RN [1]
RP SEQUENCE FROM N.A.
RA Valdez-Cruz N.A., Davila S., Licea A., Corona M., Zamudio P.Z.,
RA Garcia-Valdes J., Boyer L., Possani L.D.;
RT "Biochemical, genetic and physiological characterization of venom
RT components from two species of scorpions: Centruroides exilicauda Wood
RT and Centruroides sculpturatus Ewing.";
RL Biochimie 0:0-0(2004).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.
DR EMBL; AY649871; AAT98004.1; -.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003614; Kntol1.
DR InterPro; IPR001219; Neurotoxin.
DR Pfam; PF00537; Toxin_3; 1.
DR PRINTS; PR00285; SCORPNTOXIN.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Kntol1; 1.
DR Neurotoxin.
FT NON_TER
SQ SEQUENCE 73 AA; 8045 MW; E99165045B3A25BF CRC64;

Query Match 33.5%; Score 115; DB 2; Length 73;
Best Local Similarity 52.6%; Pred. No. 8.5e-06;
Matches 20; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 16 YLCAPLGNPDCKIKCQRKHGVDYGYCYAFQWCCEFLKD 53
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 YNCYLGRKNKYNSECKEYGVGAGYGYCYAFGCWCEGLPE 58

RESULT 13
SCX7_CENLL STANDARD; PRT; 85 AA.
AC P59865;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sodium-channel modifier neurotoxin C117 precursor.
OS Centruroides limpidus (Mexican scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthida; Buthoidea; Buthidae; Centruroides.
OX NCBI_TaxID=29941;
RN [1]
RP SEQUENCE FROM N.A.
RA Corona M., Possani L.D.;
RT "Genes and peptides from the scorpion Centruroides limpidus,
RT that recognize Na(+)-channels.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds to sodium channels and inhibits the inactivation

```

```

of the activated channels, thereby blocking neuronal transmission
(BY similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF491133; AAP49508.1; -.
DR InterPro; IPR003614; Kntol1.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; Toxin_3; 1.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Kntol1; 1.
DR Ionic channel inhibitor; Neurotoxin; Signal; Sodium channel inhibitor;
DR Toxin.
FT SIGNAL
   1 19 Potential.
FT CHAIN
   20 82 Sodium-channel modifier neurotoxin C117.
FT DISULFID
   31 82 By similarity.
FT DISULFID
   35 58 By similarity.
FT DISULFID
   44 63 By similarity.
FT DISULFID
   48 65 By similarity.
SQ SEQUENCE 85 AA; 9621 MW; 792E0B0E73618134 CRC64;

Query Match 33.5%; Score 115; DB 1; Length 85;
Best Local Similarity 52.9%; Pred. No. 9.9e-06;
Matches 18; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 16 YLCAPLGNPDCKIKCQRKHGVDYGYCYAFQWCCEFLKD 49
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 33 YICWKLGENKYCIDCKEIGAGYGYCYGFGCYCE 66

RESULT 14
Q68PG4 PRELIMINARY; PRT; 66 AA.
AC Q68PG4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Cex11 neurotoxin (Fragment).
GN Name=Cex11;
OS Centruroides exilicauda (Bark scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthida; Buthoidea; Buthidae; Centruroides.
OX NCBI_TaxID=6879;
RN [1]
RP SEQUENCE FROM N.A.
RA Valdez-Cruz N.A., Davila S., Licea A., Corona M., Zamudio P.Z.,
RA Garcia-Valdes J., Boyer L., Possani L.D.;
RT "Biochemical, genetic and physiological characterization of venom
RT components from two species of scorpions: Centruroides exilicauda Wood
RT and Centruroides sculpturatus Ewing.";
RL Biochimie 0:0-0(2004).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the alpha/beta-scorpion toxin family.
DR EMBL; AY649869; AAT98002.1; -.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003614; Kntol1.
DR InterPro; IPR001219; Neurotoxin.
DR InterPro; IPR002061; Scorpion_toxinL.
DR Pfam; PF00537; Toxin_3; 1.
DR PRINTS; PR00285; SCORPNTOXIN.
DR PRINTS; PR00284; TOXIN.
DR ProDom; PD000908; Scorpion_toxinL; 1.
DR SMART; SM00505; Kntol1; 1.
DR Neurotoxin.

```

Search completed: July 8, 2005, 15:33:03  
Job time : 179 secs

**This Page Blank (uspto)**



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	41.6	23.5	723	4	US-10-152-296-1	Sequence 1, Appli
2	40	22.6	285	1	US-08-435-040-1	Sequence 1, Appli
3	40	22.6	285	1	US-09-020-218-1	Sequence 1, Appli
4	39	22.0	2634	3	US-08-941-936-1	Sequence 1, Appli
5	38.6	21.8	615	4	US-09-902-540-6862	Sequence 6862, Ap
6	38.6	21.8	618	4	US-09-902-540-58	Sequence 58, Appli
7	38.2	21.6	1849	2	US-08-676-166A-1	Sequence 1, Appli
8	38	21.5	197	4	US-08-472-053-43	Sequence 43, Appli
9	37.2	21.0	1269	1	US-08-265-429A-4	Sequence 4, Appli
10	37.2	21.0	1269	5	PCT-US95-09069-4	Sequence 4, Appli
11	37.2	21.0	3404	1	US-08-265-429A-1	Sequence 1, Appli
12	37.2	21.0	3404	5	PCT-US95-09069-1	Sequence 1, Appli
13	37.2	21.0	3503	4	US-09-373-272-2	Sequence 2, Appli
14	36.6	20.7	394	3	US-08-866-340-22	Sequence 22, Appli
15	36.6	20.7	394	3	US-09-103-875-28	Sequence 28, Appli
16	35.8	20.2	1266	4	US-09-252-991A-1008	Sequence 1008, Ap
17	35.6	20.1	426	4	US-09-252-991A-10269	Sequence 10269, A
18	35.6	20.1	642	4	US-09-252-991A-10480	Sequence 10480, A
19	35.6	20.1	1772	4	US-09-252-991A-10690	Sequence 10690, A
20	35.4	20.0	1080	4	US-09-902-540-9563	Sequence 9563, Ap
21	35.4	20.0	14823	2	US-09-902-540-1087	Sequence 1087, Ap
22	35.2	19.9	477	2	US-08-670-186-3	Sequence 3, Appli
23	35.2	19.9	477	2	US-08-670-186-5	Sequence 5, Appli
24	35.2	19.9	950	4	US-09-640-211A-156	Sequence 156, App
25	35	19.8	50937	3	US-09-428-517-1	Sequence 1, Appli
26	34.8	19.7	1752	3	US-09-360-779-1	Sequence 1, Appli
27	34.8	19.7	1752	3	US-09-435-335-1	Sequence 1, Appli

/ GENERAL INFORMATION:  
/ APPLICANT: Hammock, Bruce D.  
/ APPLICANT: Herrmann, Rafael  
/ APPLICANT: Moskowitz, Haim  
/ TITLE OF INVENTION: Insect Control With Multiple Toxins  
/ NUMBER OF SEQUENCES: 3  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Majestic, Parsons, Siebert & Haue  
/ STREET: Four Embarcadero Center, Suite 1450  
/ CITY: San Francisco  
/ STATE: California  
/ COUNTRY: U.S.A.  
/ ZIP: 94111-4121  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/435,040  
/ FILING DATE: 08-MAY-1995  
/ CLASSIFICATION: 514  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Siebert, J. Suzanne  
/ REGISTRATION NUMBER: 28,758  
/ REFERENCE/DOCKET NUMBER: 2500.078USO  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (415) 362-5556  
/ TELEFAX: (415) 362-5418  
/ TELEX: 278638 MGPS  
/ INFORMATION FOR SEQ ID NO: 1:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 285 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: double  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: cDNA  
/ HYPOTHETICAL: NO  
/ ANTI-SENSE: NO  
/ US-08-435-040-1

Query Match 22.6%; Score 40; DB 1; Length 285;  
Best Local Similarity 61.5%; Pred. No. 0.023;  
Matches 64; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 29 ACAGCTCGGACACACTCTGCGCACCCCTGGCGGACACCGGACTGCAACA 88  
Db 86 ACATCGCGGACGACAGAACTGCGGTGTACACTGCGGCGCAACTCTTACTGCAACACCG 145

Oy 89 TCTGCCAGAGACGCGGTGCGACTACGCGTACTGCTACGCGTTC 132  
Db 146 ACTGCACCAAGACGGCGCGGACTCTGGCTACTGCCAATGGTTC 189

RESULT 3  
US-09-020-216-1  
/ Sequence 1, Application US/09020216  
/ Patent No. 6162430  
/ GENERAL INFORMATION:  
/ APPLICANT: Hammock, Bruce D.  
/ APPLICANT: Herrmann, Rafael  
/ APPLICANT: Moskowitz, Haim  
/ TITLE OF INVENTION: Insect Control With Multiple Toxins  
/ NUMBER OF SEQUENCES: 3  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Majestic, Parsons, Siebert & Haue  
/ STREET: Four Embarcadero Center, Suite 1450  
/ CITY: San Francisco  
/ STATE: California  
/ COUNTRY: U.S.A.  
/ ZIP: 94111-4121  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/020,216  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 08/435,040  
/ FILING DATE: 08-MAY-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Siebert, J. Suzanne  
/ REGISTRATION NUMBER: 28,758  
/ REFERENCE/DOCKET NUMBER: 2500.078USO  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (415) 362-5556  
/ TELEFAX: (415) 362-5418  
/ TELEX: 278638 MGPS  
/ INFORMATION FOR SEQ ID NO: 1:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 285 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: double  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: cDNA  
/ HYPOTHETICAL: NO  
/ ANTI-SENSE: NO  
/ US-09-020-216-1

Query Match 22.6%; Score 40; DB 3; Length 285;  
Best Local Similarity 61.5%; Pred. No. 0.023;  
Matches 64; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 29 ACAGCTCGGACACACTCTGCGCACCCCTGGCGGACACCGGACTGCAACA 88  
Db 86 ACATCGCGGACGACAGAACTGCGGTGTACACTGCGGCGCAACTCTTACTGCAACACCG 145

Oy 89 TCTGCCAGAGACGCGGTGCGACTACGCGTACTGCTACGCGTTC 132  
Db 146 ACTGCACCAAGACGGCGCGGACTCTGGCTACTGCCAATGGTTC 189

RESULT 4  
US-08-941-936-1  
/ Sequence 1, Application US/08941936  
/ Patent No. 6054305  
/ GENERAL INFORMATION:  
/ APPLICANT: Tatsumi, Hiroki  
/ APPLICANT: Eisaki, Naoki  
/ APPLICANT: Horiuchi, Tatsuo  
/ APPLICANT: Nagahara, Ayumu  
/ TITLE OF INVENTION: Pyruvate Orthophosphate Dikinase Gene,  
/ TITLE OF INVENTION: Recombinant DNA, and Process For Producing Pyruvate  
/ TITLE OF INVENTION: Orthophosphate Dikinase  
/ NUMBER OF SEQUENCES: 8  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: MEDLEN & CARROLL, LLP  
/ STREET: 220 Montgomery Street, Suite 2200  
/ CITY: San Francisco  
/ STATE: CA  
/ COUNTRY: US  
/ ZIP: 94104  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: Patent In Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/941,936  
/ FILING DATE: 01-OCT-1997  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Carroll, Peter G.

REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: HIRAKI-03009  
TELEPHONE: 415-705-8410  
TELEFAX: 415-397-8338  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2634 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Microbispora thermorosea  
STRAIN: IFO 14047  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2634  
US-08-941-936-1

Query Match 22.0%; Score 39; DB 3; Length 2634;  
Best Local Similarity 58.0%; Pred. No. 0.089;  
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
Qy 39 CAACACCTACTGTGGCGACCCCTGGCGGACCAACCCGGACTGCATCAAGATCTGCCAGAA 98  
Db 2475 CAAGCTCGGCATCTGCGCGAGCACGCGGGGACCCCGACTCGGTGCGGTTCCTGCCACGA 2534  
Qy 99 GCAGGCGTGCAGTACGCTACTCTACGCGCTTCCAGTGTGGTGGCGAGTTCTCTGAAGG 157  
Db 2535 GATCGGCTCGACTACGTCTCTCTGCTGCGCTTCGGCATTCGCGTGGCGCCGGCTGGAGG 2593

RESULT 5  
US-09-902-540-6862  
Sequence 6862, Application US/09902540  
Patent No. 6833447

GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wisland, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 6862  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-6862

Query Match 21.8%; Score 38.6; DB 4; Length 615;  
Best Local Similarity 52.9%; Pred. No. 0.075;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
Qy 13 GGGAACTACCGCTGGAGAGTCGCGACATCGGACACACCTACCTGTGCGCACCCCTGGGCGAAC 72  
Db 178 GGAAGCTCGCGCACCGGACAAACCCGACGATCCGGACATCGCGAGCGGCGAGTGCACATC 237  
Qy 73 CCGGACTGCATCAAGATCTGCGCAGAAGCACCGCGCTCGACTACGGCTACTGTACGCGTTC 132  
Db 238 GACTGCACCATGGCATGGCAGTTGCGGGGGTCCACTGCGCGGCGGCGCAACACGATAC 297  
Qy 133 CAGTGTGGTGGCGAGTTCTCTGAAGGACGAGAACGTCA 169  
Db 298 AACGGCTTCGGCCAGTTCTGCTGCTGGAGATGAACCCCA 334

RESULT 6

US-09-902-540-58/c  
Sequence 58, Application US/09902540  
Patent No. 6833447

GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wisland, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 58  
LENGTH: 618  
TYPE: DNA  
ORGANISM: Myxococcus xanthus  
US-09-902-540-58

Query Match 21.8%; Score 38.6; DB 4; Length 618;  
Best Local Similarity 52.9%; Pred. No. 0.075;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
Qy 13 GGGAACTACCGCTGGAGAGTCGCGACATCGGACACACCTACCTGTGCGCACCCCTGGGCGAAC 72  
Db 439 GGAAGCTCGCGCACCGGACAAACCCGACGATCCGGACATCGCGAGCGGCGAGTGCACATC 380  
Qy 73 CCGGACTGCATCAAGATCTGCGCAGAAGCACCGCGCTCGACTACGGCTACTGTACGCGTTC 132  
Db 379 GACTGCACCATGGCATGGCAGTTGCGGGGGTCCACTGCGCGGCGGCGCAACACGATAC 320  
Qy 133 CAGTGTGGTGGCGAGTTCTCTGAAGGACGAGAACGTCA 169  
Db 319 AACGGCTTCGGCCAGTTCTGCTGCGGAGATGAACCCCA 283

RESULT 7

US-08-676-166A-1  
Sequence 1, Application US/08676166A  
Patent No. 5955270  
GENERAL INFORMATION:  
APPLICANT: Radford, Alan  
APPLICANT: Parish, John H.  
TITLE OF INVENTION: EXPLOITATION OF THE CELLULOSE COMPLEX OF  
TITLE OF INVENTION: NEUROSPORA  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,166A  
FILING DATE: 15-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 1:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1849 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Neurospora crassa
; STRAIN: Oak Ridge 74A
; IMMEDIATE SOURCE:
; LIBRARY: lambda J1
; CLONE: X
; NAME/KEY: CDS
; LOCATION: join(152..832, 892..1758)
; FEATURE:
; NAME/KEY: intron
; LOCATION: 833..891
; FEATURE:
; NAME/KEY: exon
; LOCATION: <152..832
; FEATURE:
; NAME/KEY: exon
; LOCATION: 892..>1761
; PUBLICATION INFORMATION:
; AUTHORS: Radford, A
; TITLE: Cloning sequencing and homologues of the
; TITLE: CBH-1 (exocellobiohydrolase) gene of Neurospora
; TITLE: crassa
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1849
US-08-676-166A-1

Query Match 21.6%; Score 38.2; DB 2; Length 1849;
Best Local Similarity 52.9%; Pred. No. 0.14; Mismatches 0; Gaps 0;
Matches 82; Conservative 0;

QY 18 CTACCGCTCGACAGCTCGGACACACCTACCTGTGGCGACCCCTGGGCGACCAACCCGGA 77
DB 337 CTCACGAGTGTACACGGCGGCAACAGTGGCAGCGACGCTGTGCCCGATGGCAAGTC 396

QY 78 CTGCATCAAGATCTGCAGAACGACGCGCTCGACTACGGGTACTGTCTACCGGTTCCAGTG 137
DB 397 GTGCGGGCGAACTGCGCGCTGGAGCGGCGGACTACACGGGACCTACGCGGATCACCGG 456

QY 138 TTGCTGCGACTTCTGAGACGAGACGTCAGG 172
DB 457 GAGCGGCTGTCCCTCACGCTCCAGTTTCGTACCGG 491

RESULT 8
US-08-472-053-43
; Sequence 43, Application US/08472053
; Patent No. 6689356
; GENERAL INFORMATION:
; APPLICANT: ZLOTKIN, ELIAHU
; APPLICANT: MAEDA, SUSUMU
; APPLICANT: MCCUTCHEN, BILLY F.
; APPLICANT: HAMMOCK, BRUCE D.
; APPLICANT: FOWLER, ELIZABETH
; APPLICANT: BELAGAJE, RAMA M.
; TITLE OF INVENTION: RECOMBINANT BACULOVIRUSES PRODUCING
; FILE REFERENCE: UC058.1FWCP2
; CURRENT APPLICATION NUMBER: US/08/472.053
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 08/229417
; PRIOR FILING DATE: 1994-04-15
; PRIOR APPLICATION NUMBER: 07/629603
; PRIOR FILING DATE: 1990-12-19
; PRIOR APPLICATION NUMBER: 07/286087
; PRIOR FILING DATE: 1988-12-19
; NUMBER OF SEQ ID NOS: 49
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 197
; TYPE: DNA
; ORGANISM: Leivirus quinquestriatus hebraeus
; US-08-472-053-43

Query Match 21.5%; Score 38; DB 4; Length 197;
Best Local Similarity 61.4%; Pred. No. 0.079;
Matches 81; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

QY 39 CAACACCTACCTGTGGCGACCCCTGGGCGACCAACCCGAGCTGCATCAAGATCTGCCAGAA 98
DB 29 CGACGGCTGCAAGGTGGCTTGCCTGATCGGCAACGAGGGCTGCGACAAGGAGTSCAAGGC 88

QY 99 GCACGGCGTGCAGTACGGCTACTGTCTACGCGT-----TCCAGTGTGGTGGCAGTTCTT 152
DB 89 TTACGGCGGCAGCTACGGCTACTGTCTGGACCTGGGGCCCTGGCTTGTGTGGAGGGCCT 148

QY 153 GAAGGACGAGAA 164
DB 149 GCGGAGGACAA 160

RESULT 9
US-08-265-429A-4
; Sequence 4, Application US/08265429A
; Patent No. 5677151
; GENERAL INFORMATION:
; APPLICANT: Wilson, David B.
; APPLICANT: Walker, Larry P.
; APPLICANT: Zhang, Sheng
; TITLE OF INVENTION: Thermostable Cellulase From A Thermomonospora Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One Mkt Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect for Windows 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,429A
; FILING DATE: 24/06/94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 18617.0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)856-4000
; TELEFAX: 716-849-0349
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1269 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Thermomonospora fusca
; STRAIN: YX36
; CELL TYPE: bacterium
; US-08-265-429A-4

Query Match 21.0%; Score 37.2; DB 1; Length 1269;
Best Local Similarity 59.4%; Pred. No. 0.24;
Matches 63; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
```

Qy 64 GGCAGAACCCGGAGCTGCATCAAGATCTGCAGAGACGCGCTCGACTACGGCTACTGCG 123  
Db 454 GCGCGCACCCGAGCTCTCGCCCTACATGAAGCAGAACGCGCGCTACGTCACAGCGTGTGCGC 513  
Qy 124 TACCGGTTCCAGTGTGTGTGGAGTTCCTGAAGACGAGAACGTCAC 169  
Db 514 TACGCCCTCCGACAGCTGGCGGAGATCCGAACTCTACAACTACA 559

## RESULT 10

PCT-US95-09069-4

; Sequence 4, Application PC/TUS9509069

; GENERAL INFORMATION:

; APPLICANT: Wilson, David B.

; APPLICANT: Walker, Larry P.

; APPLICANT: Zhang, Sheng

; TITLE OF INVENTION: Thermostable Cellulase From A Thermomonospora Gene

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hodgson, Russ, Andrews, Woods &amp; Goodyear

; STREET: 1800 One M&amp;T Plaza

; CITY: Buffalo

; STATE: New York

; COUNTRY: United States

; ZIP: 14203-2391

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1

; SOFTWARE: Wordperfect for Windows 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/09069

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.S. Serial No. 08/265,429

; FILING DATE: 24/06/94

; ATTORNEY/AGENT INFORMATION:

; NAME: Nelson, M. Bud

; REGISTRATION NUMBER: 35,300

; REFERENCE/DOCKET NUMBER: 18617.0008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (716)856-4000

; TELEFAX: 716-849-0349

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1269 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: double-stranded

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; ORGANISM: Thermomonospora fusca

; STRAIN: YX36

; CELL TYPE: bacterium

PCT-US95-09069-4

Query Match 21.0%; Score 37.2; DB 5; Length 1269;

Best Local Similarity 59.4%; Pred. No. 0.24; Indels 0; Gaps 0;

Matches 63; Conservative 0; Mismatches 43; Mismatches 43; Indels 0; Gaps 0;

Qy 64 GGCAGAACCCGGAGCTGCATCAAGATCTGCAGAGACGCGCTCGACTACGGCTACTGCG 123  
Db 454 GCGCGCACCCGAGCTCTCGCCCTACATGAAGCAGAACGCGCGCTACGTCACAGCGTGTGCGC 513

Qy 124 TACCGGTTCCAGTGTGTGTGGAGTTCCTGAAGACGAGAACGTCAC 169  
Db 514 TACGCCCTCCGACAGCTGGCGGAGATCCGAACTCTACAACTACA 559

## RESULT 11

US-08-265-429A-1

; Sequence 1, Application US/08265429A

; Patent No. 5677151

; GENERAL INFORMATION:

; APPLICANT: Wilson, David B.

; APPLICANT: Walker, Larry P.

; APPLICANT: Zhang, Sheng

; TITLE OF INVENTION: Thermostable Cellulase From A Thermomonospora Gene

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hodgson, Russ, Andrews, Woods &amp; Goodyear

; STREET: 1800 One M&amp;T Plaza

; CITY: Buffalo

; STATE: New York

; COUNTRY: United States

; ZIP: 14203-2391

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1

; SOFTWARE: Wordperfect for Windows 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/265,429A

; FILING DATE: 24/06/94

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Nelson, M. Bud

; REGISTRATION NUMBER: 35,300

; REFERENCE/DOCKET NUMBER: 18617.0008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (716)856-4000

; TELEFAX: 716-849-0349

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3404 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: double-stranded

; TOPOLOGY: linear

; MOLECULE TYPE: Genomic DNA

; ORIGINAL SOURCE:

; ORGANISM: Thermomonospora fusca

; STRAIN: YX36

; CELL TYPE: bacterium

US-08-265-429A-1

Query Match 21.0%; Score 37.2; DB 1; Length 3404;

Best Local Similarity 59.4%; Pred. No. 0.32;

Matches 63; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 64 GGCAGAACCCGGAGCTGCATCAAGATCTGCAGAGACGCGCTCGACTACGGCTACTGCG 123  
Db 1447 GCGCGCACCCGAGCTCTCGCCCTACATGAAGCAGAACGCGCGCTACGTCACAGCGTGTGCGC 1506

Qy 124 TACCGGTTCCAGTGTGTGTGGAGTTCCTGAAGACGAGAACGTCAC 169  
Db 1507 TACGCCCTCCGACAGCTGGCGGAGATCCGAACTCTACAACTACA 1552

## RESULT 12

PCT-US95-09069-1

; Sequence 1, Application PC/TUS9509069

; GENERAL INFORMATION:

; APPLICANT: Wilson, David B.

; APPLICANT: Walker, Larry P.

; APPLICANT: Zhang, Sheng

; TITLE OF INVENTION: Thermostable Cellulase From A Thermomonospora Gene

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hodgson, Russ, Andrews, Woods &amp; Goodyear

; STREET: 1800 One M&amp;T Plaza

; CITY: Buffalo

; STATE: New York

; COUNTRY: United States

; ZIP: 14203-2391

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect for Windows 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09069
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial No. 08/265,429
; FILING DATE: 24/06/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 18617.0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)856-4000
; TELEFAX: 716-849-0349
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3404 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Thermomonospora fusca
; STRAIN: YX36
; CELL TYPE: bacterium
; PCT-US95-09069-1

Query Match 21.0%; Score 37.2; DB 5; Length 3404;
Best Local Similarity 59.4%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 43;

QY 64 GCGGCAACCCGAGCTGATCAAGATCTGCCAGAACGCGCGTCCGACTACGGTACTGC 123
Db 1447 GCGGCAACCCGAGCTGCTGCGCTTACATGAAGCAGACGCGGTACGTCACGGTGTGGC 1506
QY 124 TAGCGGTTCCAGTGTGTGTCGGAGTTCTCTGAAGGACGAGAACGTCA 169
Db 1507 TAGCGGCTCCGCAAGCTGGCGGAGATCCCGAACGTCCTACAACATACA 1552

RESULT 13
US-09-373-272-2
; Sequence 2, Application US/09373272
; Patent No. 6818803
; GENERAL INFORMATION:
; APPLICANT: Austin-Phillips, Sandra
; APPLICANT: Burgess, Richard D
; APPLICANT: German, Thomas L
; APPLICANT: Ziegelhofer, Thomas
; TITLE OF INVENTION: Transgenic Plants as an Alternative Source of
; EARLIER INVENTION: Lignocellulosic-Degrading Enzymes
; FILE REFERENCE: Transgenic Plants Expressing Cellulase
; CURRENT APPLICATION NUMBER: US/09/373,272
; CURRENT FILING DATE: 1999-08-12
; EARLIER APPLICATION NUMBER: 08/883,495
; EARLIER FILING DATE: 1997-06-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3503
; TYPE: DNA
; ORGANISM: Thermomonospora fusca
; US-09-373-272-2

Query Match 21.0%; Score 37.2; DB 4; Length 3503;
Best Local Similarity 59.4%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 43;

QY 64 GCGGCAACCCGAGCTGATCAAGATCTGCCAGAACGCGCGTCCGACTACGGTACTGC 123
Db 1447 GCGGCAACCCGAGCTGCTGCGCTTACATGAAGCAGACGCGGTACGTCACGGTGTGGC 1506
QY 124 TAGCGGTTCCAGTGTGTGTCGGAGTTCTCTGAAGGACGAGAACGTCA 169
Db 1507 TAGCGGCTCCGCAAGCTGGCGGAGATCCCGAACGTCCTACAACATACA 1552

RESULT 14
US-08-866-340-22
; Sequence 22, Application US/08866340
; Patent No. 6020318
; GENERAL INFORMATION:
; APPLICANT: Szyf, Moshe
; APPLICANT: Bigey, Pascal
; APPLICANT: Ramchandani, Shyam
; TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC
; TITLE OF INVENTION: SEQUENCES AND ANTISENSE OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,340
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Keown, Wayne A.
; REGISTRATION NUMBER: 33,923
; REFERENCE/DOCKET NUMBER: 106.101.187
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-5000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-866-340-22

Query Match 20.7%; Score 36.6; DB 3; Length 394;
Best Local Similarity 53.1%; Pred. No. 0.25;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 30 CAGCTCGGACACACCTACTGTGCGCACCCCTGGGGGACACCCCGACTGCTCAAGAT 89
Db 37 CAGGCTTGAAGACACCCCAAGTCCACTCCAGGAGTACCACGACATCAACCTGCT 96
QY 90 CTGCCAGAACGACGGCGTGACTACGGCTACTGCTACGGGTTCCAGTGTGGTGGAGTT 149
Db 97 CTACTGGAGGACGAGAGGCGCGTGTGGACTTCAAGGTTGTGCAGGCGCGCTGCACGT 156
QY 150 CTTGAAGGACGAGAACGTCGAAGGTGTG 176
Db 157 GGAGTATGGGAGGACCTGCCCGAGTG 183

RESULT 15
US-09-103-875-28
```





This Page blank (uspio,

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:25:34 ; Search time 255.059 Seconds  
(without alignments)  
4108.047 Million cell updates/sec

Title: US-10-617-978-17\_COPY\_73\_249  
Perfect score: 177  
Sequence: 1 gccagcgtgccagggaacta.....acgagaacgtcaagggtgtga 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870567 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2002bs.\*
- 8: Geneseq2003as.\*
- 9: Geneseq2003bs.\*
- 10: Geneseq2003cs.\*
- 11: Geneseq2003ds.\*
- 12: Geneseq2004as.\*
- 13: Geneseq2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	23.5	675	6	AAD46281 Discosoma
2	41.6	23.5	678	6	AAD46278 Discosoma
3	41.6	23.5	678	6	AAD46282 Discosoma
4	41.6	23.5	678	6	AAD28208 Discosoma
5	41.6	23.5	678	6	AAD28207 Discosoma
6	41.6	23.5	678	6	AAD28209 Discosoma
7	41.6	23.5	678	10	ADC24128 Discosoma
8	41.6	23.5	678	11	ADL46207 Discosoma
9	41.6	23.5	681	10	ADC24134 Discosoma
10	41.6	23.5	681	11	ADL46205 Human cod
11	41.6	23.5	681	11	ADL46225 Human cod
12	41.6	23.5	695	3	AA448743 Humanised
13	41.6	23.5	704	11	ADN33980 Mutant-ty
14	41.6	23.5	723	8	AB222476 Mammalian
15	41.6	23.5	1638	10	ADL18131 RFP:PS(NI
16	41.6	23.5	1647	10	ADL18155 RFP:PS(HI
17	41.6	23.5	2721	13	ADS75466 Fibrohexa
18	41.6	23.5	2772	13	ADS75468 Fibrohexa
19	41.6	23.5	4488	13	ADQ91220 Fusion co
20	41.6	23.5	4692	6	AAL47954 Modified

21	41.6	23.5	4692	10	ACC44640
22	41.6	23.5	5436	4	AD10003
23	41.6	23.5	6706	12	ADL35208
24	41.6	23.5	6893	10	ADL24111
25	41.6	23.5	6985	10	ACA55359
26	41.6	23.5	7495	13	ADR30815
27	41.6	23.5	7508	13	ADR30814
28	41.6	23.5	7616	13	ADQ91222
29	41.6	23.5	7910	4	ADQ9979
30	41.6	23.5	7927	12	ADL35211
31	41.6	23.5	9320	6	ABS56664
32	41.6	23.5	9658	12	ADI04104
33	41.6	23.5	9678	12	ADI04103
34	41.6	23.5	9731	11	ADM82791
35	41.6	23.5	9782	11	ADM82792
36	41.6	23.5	10263	12	ADI04102
37	41	23.2	726	10	ADH00962
38	40.4	22.8	263	7	ADS66256
39	40.4	22.8	714	12	ADJ44616
40	40	22.6	285	2	AA742608
41	40	22.6	678	10	ADC24132
42	40	22.6	678	11	ADL46282
43	40	22.6	678	11	ADL46211
44	39.8	22.5	3128	5	AA578582
45	39.8	22.5	5092	4	AAI59458

ALIGNMENTS

RESULT 1

AAD46281

ID AAD46281 standard; DNA; 675 BP.

AC AAD46281;

XX 27-DEC-2002 (first entry)

XX Discosoma sp. drFP583 (NFP-6) mutant DNA, E57-NA.

KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreener; second messenger detector; drFP583 protein; NFP-6; gene;  
KW mutant; ds.  
XX Discosoma sp.  
OS Synthetic.

XX Key

XX CDS

XX Location/Qualifiers

XX 1..675

XX /\*tag= a

XX /product= "drFP583 mutant protein, E57-NA"

XX /note= "No stop codon"

XX /partial

XX WO200268459-A2.

XX PD

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX PF

XX 21-FEB-2001; 2001US-0270983P.

XX PR

XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukanov S, Lukanov K, Yanushevich Y, Savitsky A, Fradkov A;

XX WPI; 2002-691654/74.

XX P-PSDB; AAE28836.

XX







```
SQ Sequence 678 BP; 145 A; 226 C; 203 G; 104 T; 0 U; 0 Other;
Query Match 23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.44;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 2 CCGACGTCGCCAGGAACTACCGCTGGACAGCTCGGACACACCTACTGTGGCACCCC 61
DB 206 CCAAGGTGTACGTGAAGACACCCCGCGACATCCCGGACTACAAGAGTGTCTTCCCG 265
QY 62 TGGCGGACAAACCCCGACTGCATCAAGATCTGCCAGAAAGCAGCGGTGCACTACCGGTACT 121
DB 266 AGGCTTCAAGTGGAGCGCGTGTGATCACTTCGAGGACGCGGCTGGACCGTGACCC 325
QY 122 GCTACGCGTTCACGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
DB 326 AGGACTCCTCCTGCAGGACGCGTCTTCTATCTACAAGGTGAAGTTCA 373

RESULT 7
ADC24128
XX ADL46207 standard; DNA; 678 BP.
XX AC ADC24128;
XX DT 18-DEC-2003 (first entry)
XX DE Discosoma red fluorescent protein variant fast T1 DNA.
XX KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;
XX KW fluorescent protein variant; transcription induction detection;
XX KW fluorescence energy resonance transfer; FRET; protein kinase;
XX KW protein phosphatase; ion indicator; ds; mutant; fast T1.
XX OS Synthetic.
XX OS Discosoma.
XX PN US2003059835-A1.
XX PD 27-MAR-2003.
XX PF 10-APR-2002; 2002US-00121258.
XX PR 26-FEB-2001; 2001US-00794308.
XX PR 24-MAY-2001; 2001US-00866538.
XX PA (TSIE/) TSIE R Y.
XX PA (CAMP/) CAMPBELL R E.
XX PI Tsien RY, Campbell RE;
XX PI WPI; 2003-743764/70.
XX DR P-PSDB; ADC24129.
XX PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein
XX PT variant having a reduced propensity to oligomerize, useful for detecting
XX PT transcriptional activity.
XX PS Example 1; SEQ ID NO 3; 67pp; English.
XX CC The invention describes a polynucleotide sequence (I) encoding a
XX CC Discosoma red fluorescent protein (DsRed) variant having a reduced
XX CC propensity to oligomerize, comprising amino acid substitutions at the AB
XX CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225
XX CC amino acids, given in the specification, where the substitutions result
XX CC in reduced propensity of the DsRed variant to form tetramers. (I) is
XX CC useful for detecting transcriptional activity by providing a host cell
XX CC containing a vector which comprises (I) operatively linked to an
XX CC expression control sequence, and an unit to assay the variant fluorescent
XX CC protein fluorescence, and assaying fluorescence of the variant
XX CC fluorescent protein produced by (VII), where variant fluorescent protein
XX CC fluorescence is indicative of transcriptional activity. A polynucleotide
XX CC encoding a fusion protein is useful for the analysis of in vivo
```

```
CC localisation or trafficking of a polypeptide of interest. A polypeptide
CC marker is useful as markers to identify the location and amount of a
CC target protein produced, where the target protein is fused to the marker,
CC as a complement to or alternative for the green fluorescent protein or
CC its spectral variant, for detecting induction of transcriptions, in
CC applications involving fluorescence energy resonance transfer (FRET),
CC which detects events as the function of the movement of fluorescent
CC donors and acceptors towards or away from each other, for making
CC fluorescent sensors for protein kinase and phosphatase activities or
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the
CC presence of a molecule in a sample, for identifying a specific
CC interaction of a first and second molecule, for determining whether a
CC sample contains an enzyme or for determining the pH of the sample. (I) is
CC useful for identifying a region or condition that regulates the activity
CC of an expression control sequence. This sequence encodes Discosoma red
CC fluorescent protein variant fast T1.
SQ Sequence 678 BP; 141 A; 225 C; 205 G; 107 T; 0 U; 0 Other;
Query Match 23.5%; Score 41.6; DB 10; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.44;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 2 CCGACGTCGCCAGGAACTACCGCTGGACAGCTCGGACACACCTACTGTGGCACCCC 61
DB 206 CCAAGGTGTACGTGAAGACACCCCGCGACATCCCGGACTACAAGAGTGTCTTCCCG 265
QY 62 TGGCGGACAAACCCCGACTGCATCAAGATCTGCCAGAAAGCAGCGGTGCACTACCGGTACT 121
DB 266 AGGCTTCAAGTGGAGCGCGTGTGATCACTTCGAGGACGCGGCTGGACCGTGACCC 325
QY 122 GCTACGCGTTCACGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
DB 326 AGGACTCCTCCTGCAGGACGCGTCTTCTATCTACAAGGTGAAGTTCA 373

RESULT 8
ADL46207
XX ID ADL46207 standard; DNA; 678 BP.
XX AC ADL46207;
XX DT 20-MAY-2004 (first entry)
XX DE Discosoma red fluorescent protein (DsRed) variant T1 coding sequence.
XX KW ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;
XX KW oligomerization; tetramerization; immunoassay; hybridization assay.
XX OS Discosoma sp.
XX OS Key Location/Qualifiers
XX FH CDS 1..678
XX FT /*tag= a
XX FT /product= "DsRed variant T1 protein"
XX PN WO2003086446-A1.
XX PD 23-OCT-2003.
XX PF 09-APR-2003; 2003WO-US010879.
XX PR 10-APR-2002; 2002US-00121258.
XX PR 29-JUL-2002; 2002US-00209208.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Tsien RY, Campbell RE, Baird GS;
XX DR WPI; 2003-845265/78.
XX DR P-PSDB; ADL46206.
XX PT New monomeric and dimeric Anthozoan fluorescent protein variants with
```





```
PD 23-OCT-2003.
XX 09-APR-2003; 2003WO-US010879.
XX 10-APR-2002; 2002US-00121258.
XX 29-JUL-2002; 2002US-00209208.
XX (REGC ) UNIV CALIFORNIA.
XX Tsien RY, Campbell RE, Baird GS;
XX WPI; 2003-845265/78.
XX New monomeric and dimeric Anthozoan fluorescent protein variants with
XX reduced propensity to oligomerize, and encoding polynucleotides, useful
XX in molecular biology, e.g. in immunoassays or in tracking protein
XX movement in cells.
XX Disclosure; SEQ ID NO 3; 166pp; English.
XX The invention relates to a polynucleotide sequence encoding a Discosoma
XX red fluorescent protein (DsRed) variant having a reduced propensity to
XX oligomerize. The protein variant comprises one or more amino acid
XX substitutions at the AB and/or AC interface(s) of the wild-type DsRed
XX sequence, where the substitutions result in reduced propensity of the
XX DsRed variant to form tetramers and where the variant displays detectable
XX fluorescence of at least one red wavelength. The composition and methods
XX are useful in producing red fluorescent proteins having reduced
XX propensity for oligomerization, especially tetramerization. The protein
XX may be used in molecular biology and in other scientific applications,
XX such as in immunoassays or hybridization assays, or in tracking the
XX movement of proteins in cells. This sequence corresponds to the DsRed
XX coding sequence with codons optimised for human expression.
XX Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 11; Length 681;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
Oy 2 CCGAGCTGCCAGGGAATACCGCTGGACAGCTCGGACACACCTACCTGTGGCACC 61
Db 209 CCAAGGTGTAGTGAAGCACCCTCCGACATCCCGACTACAGAGCTGTCTTCCCG 268
Oy 62 TGGCGCAACACCCGAGTGCATCAAGATCTGCCAGACGACGGCTGCGCTACT 121
Db 269 AGGCTTCAAGTGGAGCGCTGATGAATTCGAGGACGGCGGTGTGACCGTGACCC 328
Oy 122 GCTACGGCTTCCAGTGTGGTGCGAGTTCCTGGAAGGACGAGAAGCTCA 169
Db 329 AGGACTCTCCTCGAGGACGGCTGCTTCACTACAGGTGAAGTTCA 376
XX
RESULT 11
ADL46225
ID ADL46225 standard; DNA; 681 BP.
XX AC ADL46225;
XX AC
XX 20-MAY-2004 (first entry)
XX DE Human codon optimised Discosoma red fluorescent protein (DsRed) DNA.
XX ds; Gene; red fluorescent protein; DsRed; fluorescence; red wavelength;
XX oligomerization; tetramerization; immunoassay; hybridization assay.
XX OS Discosoma sp.
XX WO2003086446-A1.
XX PN
XX 23-OCT-2003.
XX 09-APR-2003; 2003WO-US010879.
XX
XX Query Match 23.5%; Score 41.6; DB 11; Length 681;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
Oy 2 CCGAGCTGCCAGGGAATACCGCTGGACAGCTCGGACACACCTACCTGTGGCACC 61
Db 209 CCAAGGTGTAGTGAAGCACCCTCCGACATCCCGACTACAGAGCTGTCTTCCCG 268
Oy 62 TGGCGCAACACCCGAGTGCATCAAGATCTGCCAGACGACGGCTGCGCTACT 121
Db 269 AGGCTTCAAGTGGAGCGCTGATGAATTCGAGGACGGCGGTGTGACCGTGACCC 328
Oy 122 GCTACGGCTTCCAGTGTGGTGCGAGTTCCTGGAAGGACGAGAAGCTCA 169
Db 329 AGGACTCTCCTCGAGGACGGCTGCTTCACTACAGGTGAAGTTCA 376
XX
RESULT 12
AAA48743
ID AAA48743 standard; cDNA; 695 BP.
XX AC AAA48743;
XX AC
XX 19-SEP-2000 (first entry)
XX DT
XX DE Humanised Discosoma sp. "red" novel fluorescent protein drFP583 cDNA.
XX Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;
XX fluorescent labeling; ss.
XX OS Discosoma sp; "red".
XX OS Synthetic.
XX WO2000034326-A1.
XX PN
XX 15-JUN-2000.
XX PD
XX 10-DEC-1999; 99WO-US029473.
XX PF
XX 11-DEC-1998; 98US-00210330.
XX PR
```

```
XX 10-APR-2002; 2002US-00121258.
XX 29-JUL-2002; 2002US-00209208.
XX (REGC ) UNIV CALIFORNIA.
XX Tsien RY, Campbell RE, Baird GS;
XX WPI; 2003-845265/78.
XX New monomeric and dimeric Anthozoan fluorescent protein variants with
XX reduced propensity to oligomerize, and encoding polynucleotides, useful
XX in molecular biology, e.g. in immunoassays or in tracking protein
XX movement in cells.
XX Disclosure; SEQ ID NO 23; 166pp; English.
XX The invention relates to a polynucleotide sequence encoding a Discosoma
XX red fluorescent protein (DsRed) variant having a reduced propensity to
XX oligomerize. The protein variant comprises one or more amino acid
XX substitutions at the AB and/or AC interface(s) of the wild-type DsRed
XX sequence, where the substitutions result in reduced propensity of the
XX DsRed variant to form tetramers and where the variant displays detectable
XX fluorescence of at least one red wavelength. The composition and methods
XX are useful in producing red fluorescent proteins having reduced
XX propensity for oligomerization, especially tetramerization. The protein
XX may be used in molecular biology and in other scientific applications,
XX such as in immunoassays or hybridization assays, or in tracking the
XX movement of proteins in cells. This sequence corresponds to the DsRed
XX coding sequence with codons optimised for human expression.
XX Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 11; Length 681;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
Oy 2 CCGAGCTGCCAGGGAATACCGCTGGACAGCTCGGACACACCTACCTGTGGCACC 61
Db 209 CCAAGGTGTAGTGAAGCACCCTCCGACATCCCGACTACAGAGCTGTCTTCCCG 268
Oy 62 TGGCGCAACACCCGAGTGCATCAAGATCTGCCAGACGACGGCTGCGCTACT 121
Db 269 AGGCTTCAAGTGGAGCGCTGATGAATTCGAGGACGGCGGTGTGACCGTGACCC 328
Oy 122 GCTACGGCTTCCAGTGTGGTGCGAGTTCCTGGAAGGACGAGAAGCTCA 169
Db 329 AGGACTCTCCTCGAGGACGGCTGCTTCACTACAGGTGAAGTTCA 376
XX
RESULT 12
AAA48743
ID AAA48743 standard; cDNA; 695 BP.
XX AC AAA48743;
XX AC
XX 19-SEP-2000 (first entry)
XX DT
XX DE Humanised Discosoma sp. "red" novel fluorescent protein drFP583 cDNA.
XX Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;
XX fluorescent labeling; ss.
XX OS Discosoma sp; "red".
XX OS Synthetic.
XX WO2000034326-A1.
XX PN
XX 15-JUN-2000.
XX PD
XX 10-DEC-1999; 99WO-US029473.
XX PF
XX 11-DEC-1998; 98US-00210330.
XX PR
```

```
PR 14-OCT-1999; 99US-00418529.
XX (CLON-) CLONTECH LAB INC.
PA Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;
XX Ding L;
XX WPI; 2000-423381/36.
XX Novel fluorescent protein from non-bioluminescent Discosoma sp. red,
XX useful for fluorescent labeling and as markers.
XX Claim 6; Page 75-76; 86pp; English.
XX The present sequence is humanised drFP583 cDNA. drFP583 is a full-length
XX cDNA encoding a novel fluorescent protein (nFP) from Discosoma sp. "red",
XX a non-bioluminescent species of the Class Anthozoa. The wild-type drFP583
XX nucleotide sequence was altered to optimise the codons for expression of
XX the fluorescent protein in mammalian cells. Fluorescent proteins can be
XX used in fluorescent labeling, a useful tool for marking a protein, cell
XX or organism of interest. Unlike other markers used in protein labeling,
XX such as beta-galactosidase and luciferase, fluorescent proteins do not
XX require an exogenous cofactor or substrate. Methods involving fluorescent
XX proteins are also less laborious and less difficult to control than the
XX traditional methods of fluorescent labeling, where a protein of interest
XX is purified and then covalently conjugated to a fluorophore derivative.
XX Novel fluorescent proteins isolated from species of the Class Anthozoa
XX can be used as markers for gene expression and protein localization
XX studies, and in fluorescence resonance energy transfer (FRET) reactions.
XX They may have improved properties and better suitability for larger
XX excitations compared to prior art fluorescent proteins such as green
XX fluorescent protein
XX
XX Sequence 695 BP; 149 A; 228 C; 209 G; 109 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 3; Length 695;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
XX QY 2 CCGACGTGCCAGGAACTACCGGTGGACAGCTGGGACACACCTACTCTGGCGCACCCC 61
XX Db 215 CCAAGGTGTACGTGAAGCACCCTCCGACATCCCGGACTCAAGAAGCTGTCTTCCCG 274
XX
XX QY 62 TGGCGACAACCCGGACTGCATCAAGATCTGCCAGAAGCACGGCGTGCAGTACGGCTACT 121
XX Db 275 AGGCTTCAGTGGAGCGGTGATGATCTCGAGGACGGCGGTGTGACCGTGACCC 334
XX
XX QY 122 GCTACGGCTTCCAGTTGTGTGGAGTTCTCTGAAGGACGAGAGCTCA 169
XX Db 335 AGGACTCTCCCTGCAGGCGGTGCTTTCATCTACAAGGTGAAGTTCA 382
XX
XX RESULT 13
XX ADN33980
XX ID ADN33980 standard; DNA; 704 BP.
XX
XX AC ADN33980;
XX
XX DT 01-JUL-2004 (first entry)
XX
XX DE Mutant-type D8RED encoding sequence.
XX
XX KW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type D8RED;
XX mutant; mteuin; ds.
XX
XX OS Discosoma sp.
XX
XX PN WO2003054158-A2.
XX
XX PD 03-JUL-2003.
XX
XX PF 18-DEC-2002; 2002WO-US040539.
XX
```

```
PR 19-DEC-2001; 2001US-0341723P.
XX (UYCH-) UNIV CHICAGO.
PA Bevis B, Glick B;
XX WPI; 2003-569236/53.
XX Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent
XX mutant of a Cnidarian chromo- or fluorescent protein or its mutant,
XX useful for applications involving chromo- or fluorescent proteins.
XX Claim 7; SEQ ID NO 3; 65pp; English.
XX The present invention relates to nucleic acid that encodes a rapidly
XX maturing chromo or fluorescent mutant of a Cnidarian chromo- or
XX fluorescent protein or its mutant. The protein is useful in applications
XX involving nucleic acid encoding a chromo- or fluorescent protein and is
XX useful for producing a chromo and/or fluorescent protein which involves
XX growing the cell, whereby the protein is expressed, and isolating the
XX protein substantially free of other proteins. The protein is useful in
XX applications involving chromo- or fluorescent protein and is useful as
XX PCR primers, hybridization probes, etc. The expression cassettes are
XX useful for synthesizing related proteins. The chromoproteins are useful
XX as coloring agents which are capable of imparting color or pigment to a
XX particular composition of matter e.g. food compositions, pharmaceuticals,
XX cosmetics, living organisms, e.g., animals and plants. The chromoproteins
XX may also find use as labels in analyte detection assays, e.g. assays for
XX biological analytes of interest and as selectable markers in recombinant
XX DNA applications, e.g. the production of transgenic cells and organisms.
XX The fluorescent proteins find use in a variety of different applications,
XX e.g. in fluorescence resonance energy transfer (FRET) applications, as
XX biosensors in prokaryotic and eukaryotic cells, in applications involving
XX the automated screening of arrays of cells expressing fluorescent
XX reporting groups by using microscopic imaging and electronic analysis, as
XX second messenger detectors, and in fluorescence activated cell sorting
XX applications and as in vivo marker in animals. The fluorescent proteins
XX also find use in protease cleavage assays. The proteins can also be used
XX as assays to determine the phospholipid composition in biological
XX membranes and as a fluorescent timer. The present sequence represents the
XX mutant-type D8RED encoding sequence.
XX
XX SQ Sequence 704 BP; 145 A; 237 C; 212 G; 110 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 11; Length 704;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
XX QY 2 CCGACGTGCCAGGAACTACCGGTGGACAGCTGGGACACACCTACTCTGTGCGCACCCC 61
XX Db 225 CCAAGGTGTACGTGAAGCACCCTCCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 284
XX
XX QY 62 TGGCGACAACCCGGACTGCATCAAGATCTGCCAGAAGCACGGCGTGCAGTACGGCTACT 121
XX Db 285 AGGCTTCAGTGGAGCGCGGTGATGAACTTCGAGGACGGCGGTGGTGACCGTGACCC 344
XX
XX QY 122 GCTACCGCTTCCAGTTGTGTGGAGTTCTCTGAAGGACGAGAGCTCA 169
XX Db 345 AGGACTCTCCCTGCAGGCGGTCTCTTATCTACAAGGTGAAGTTCA 392
XX
XX RESULT 14
XX ABZ22476
XX ID ABZ22476 standard; DNA; 723 BP.
XX
XX AC ABZ22476;
XX
XX DT 25-MAR-2003 (first entry)
XX
XX DE Mammalian codon optimised Discosoma red fluorescent protein DNA.
XX
XX KW Mammalian codon optimised Discosoma red fluorescent protein; Discosoma;
XX red fluorescent protein; directed evolution; gene; ds.
```

```
XX Discosoma sp.
OS Mammalia.
OS Synthetic.
XX Key Location/Qualifiers
XX CDS 1..723
XX FT /*tag= a
XX FT /product= "Mammalian codon optimised Discosoma red
XX FT fluorescent protein"
XX FT /transl_except= (pos:616..618,aa:His)
XX XX
XX W0200294992-A2.
XX XX
XX 28-NOV-2002.
XX XX
XX 20-MAY-2002; 2002WO-US015968.
XX XX
XX 18-MAY-2001; 2001US-0291871P.
XX XX
XX (RIGE-) RIGEL PHARM INC.
XX PA
XX Peelle B;
XX PI
XX WPI; 2003-120798/11.
XX DR
XX P-PSDB; ABP56678.
XX XX
XX New Discosoma red fluorescent protein, useful for functional screens as a
XX reporter for gene transcription, for target characterization and
XX localization of fusion proteins, or for scaffolds for protein and peptide
XX libraries.
XX XX
XX Example 1; Fig 1; 22pp; English.
XX XX
XX The present invention describes an isolated Discosoma red fluorescent
XX protein (I) comprising a 241 residue amino acid sequence (see ABP56678,
XX S1), with one or more point mutations at amino acid position N24, F125,
XX K164, or M183. Also described: (1) a fusion protein comprising (I); (2)
XX an isolated nucleic acid encoding (1); (3) a vector comprising the
XX nucleic acid of (2); (4) a host cell comprising the vector; (5) a
XX retroviral cDNA expression library comprising the nucleic acid of (2), or
XX encoding (1); and (6) methods of making a fluorescent variant. The
XX Discosoma red fluorescent proteins are useful for functional screens as a
XX reporter for gene transcription (e.g. as a fusion protein), for target
XX characterisation and localisation of fusion proteins, or for scaffolds
XX for protein and peptide libraries. The fluorescent proteins can also be
XX used as selectable markers or reporter molecules for a variety of
XX bioassays, including methods that use fluorescence activated cell sorting
XX (FACS) as a selection mechanism. The method of directed protein evolution
XX is useful for obtaining improved variants of red fluorescent protein. The
XX variants of Discosoma red fluorescent protein have greatly improved
XX brightness, expression, and/or folding kinetics as compared to wild type
XX or a codon optimised variant. The present sequence encodes a mammalian
XX codon optimised Discosoma red fluorescent protein, which is used in an
XX example from the present invention
XX XX
XX Sequence 723 BP; 161 A; 233 C; 221 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 8; Length 723;
XX Best Local Similarity 53.0%; Pred. No. 0.44;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
XX
XX 2 CCAGCGTCCAGGGAACCTACCGCTGGACACGACCTCGGACCAACACCTACCTGTGCGACCCC 61
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 209 CCAAGTGTAGCTGAGACACCCCGCCGACATCCCGACTACAGAGCTGCTCTCCCG 268
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 62 TGGGCGCAACCCCGACTGTGATCAAGATTCCTCCCAAGACGCGGTGCTACGCTACT 121
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 269 AGGCGTTCAAGTGGGAGCGCGTGTGATGAACCTTCGAGGACGGCGGTGTGACCGTACCC 328
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 122 GCTACGCGTTCAGTGTGGTGGAGTTCCTGAGGACGAGACGTCA 169
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 329 AGGACTCCTCCCTGCAGGACGGCTGCTTCTATCAAGGTGAAGTTCA 376
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 15
ADL18131
XX ADL18131 standard; DNA; 1638 BP.
XX AC ADL18131;
XX XX
XX 06-MAY-2004 (first entry)
XX DT
XX DE RFP:PS (N1a protease):AtOEP7:GFP fusion protein DNA SEQ ID NO:51.
XX XX
XX chimeric protein; signal protein; trafficking signal targeting;
XX KW proteolytic cleavage site; protease; protease inhibitor; gene; ds.
XX OS
XX Arabidopsis thaliana.
XX OS
XX Tobacco vein mottling virus.
XX OS
XX Synthetic.
XX XX
XX W02003014381-A1.
XX PN
XX 20-FEB-2003.
XX PD
XX 08-AUG-2002; 2002WO-KR001515.
XX PF
XX 10-AUG-2001; 2001KR-00048123.
XX PR
XX (AHRA-) AHRAH BIOSYSTEMS INC.
XX PA
XX Hwang I, Kim DH, Lee YJ;
XX PI
XX WPI; 2003-256596/25.
XX DR
XX P-PSDB; ADL18132.
XX DR
XX
XX New chimeric protein, useful for detecting protease inhibitors inside the
XX cell or tissue.
XX
XX Example 2; SEQ ID NO 51; 214pp; English.
XX
XX The present invention describes a chimeric protein comprising at least
XX one signal protein that has a trafficking signal targeting to a
XX subcellular organelle and at least one proteolytic cleavage site for a
XX protease. The chimeric protein is constructed, so that: (a) the
XX trafficking signals of all the signal proteins are inactivated by linking
XX the proteolytic site or a signal masking protein through the proteolytic
XX site to the N- or C-terminus of the signal proteins, and so the chimeric
XX protein is present in cytosol; (b) the trafficking signal of at least one
XX signal protein is activated when the proteolytic cleavage site is cleaved
XX by the protease, and as a result at least one fragment protein that
XX includes the activated signal protein is a transported to a subcellular
XX organelle; and (c) the chimeric protein is labelled with at least one
XX fluorescent label signal in the cell is altered depending on the cleavage
XX by the protease. Also described: (1) a recombinant gene comprising a
XX nucleic acid sequence encoding the chimeric protein which is constructed
XX to express the chimeric protein in a cell; (2) a cell transformed with
XX the recombinant gene or vector; (3) analysing the activity of a protease
XX in vivo; (4) screening protease inhibitors in vivo; (5) a system for
XX detecting a protease inside a cell; (6) a nucleic acid comprising the
XX sequence encoding the chimeric protein for detecting protease activity in
XX a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting
XX a protease inside a cell comprising the chimeric protein or the vector;
XX (9) detecting a protease inside a cell or tissue; and (10) detecting a
XX protease inhibitor in vivo. The chimeric protein is useful for detecting
XX protease inhibitors inside the cell or tissue. The present sequence
XX encodes a fusion protein, which is used in the exemplification of the
XX present invention.
XX
XX Sequence 1638 BP; 445 A; 450 C; 419 G; 324 T; 0 U; 0 Other;
XX
XX Query Match 23.5%; Score 41.6; DB 10; Length 1638;
XX Best Local Similarity 53.0%; Pred. No. 0.48;
XX Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

Qy	2	CCGACGTGCCAGGGAAC	TACCCGCTGGACAGCTGG	AACAACCTACCTGTGCGACCCC	61
Db	209	CCAAGGTGTACGTGAAG	CAACCCCGCCGACATCCC	CGACTACNAGAAGCTGCTTCC	268
Qy	62	TGGCGGACAACCCGGAC	TGCATCAGATCTGCCAG	AAGCAGCGGCTCGACTACGGCTACT	121
Db	269	AGGGCTTCAAAGTGGG	AGCGGTGATGAAC	TCGAGGACGGCGGTGGTGACCGTGACCC	328
Qy	122	GCTACGGGTTCCAGTGT	TGGTGCGAGTTCTCTGA	AGGACGAGAGCTCA	169
Db	329	AGGACTCTCCTCCCTG	CAGGCGGTGCTTCA	TCTACAAGGTGAAGTTCA	376

Search completed: July 30, 2005, 18:28:48  
Job time : 259,059 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:39:09 ; Search time 1262.87 Seconds  
(without alignments)  
6791.361 Million cell updates/sec

Title: US-10-617-978-17\_COPY\_73\_249

Perfect score: 177

Sequence: 1 gccagctgccagggaacta.....acgagaactcaaggtgtga 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.8	25.9	354	3 AF151798	AF151798 Mesobuthus
2	45.8	25.9	469	3 AY282464	AY282464 Mesobuthu
3	43.8	24.7	299925	1 AP005043	AP005043 Streptomy
4	43.6	24.6	110000	1 AP006840	Continuation (33 O
5	41.6	23.5	316	3 AF491132	AF491132 Centruoi
6	41.6	23.5	675	6 AX824731	AX824731 Sequence
7	41.6	23.5	678	6 AX370404	AX370404 Sequence
8	41.6	23.5	678	6 AX370406	AX370406 Sequence
9	41.6	23.5	678	6 AX370408	AX370408 Sequence
10	41.6	23.5	678	6 AX824725	AX824725 Sequence
11	41.6	23.5	678	6 AX824732	AX824732 Sequence
12	41.6	23.5	723	6 AR527331	AR527331 Sequence
13	41.6	23.5	1050	6 AX666133	AX666133 Sequence
14	41.6	23.5	2721	6 CQ882115	CQ882115 Sequence
15	41.6	23.5	2772	6 CQ882117	CQ882117 Sequence
16	41.6	23.5	4488	6 CQ849509	CQ849509 Sequence
17	41.6	23.5	4692	6 AX463702	AX463702 Sequence
18	41.6	23.5	6423	12 AY613997	AY613997 Cloning v
19	41.6	23.5	6893	6 AX823860	AX823860 Sequence

20	41.6	23.5	7616	6 CQ849511	CQ849511 Sequence
21	41.6	23.5	8811	12 AY569779	AY569779 Cloning v
22	41.6	23.5	9320	6 AX663075	AX663075 Sequence
23	41.6	23.5	10141	12 AY342347	AY342347 Red H-Pel
24	41.6	23.5	10276	12 AY342348	AY342348 Red H-Sti
25	41.6	23.5	10481	12 AY490568	AY490568 UAS-Red S
26	41.6	23.5	12404	12 AY569780	AY569780 Cloning v
27	41.4	23.4	3093	9 AK122759	AK122759 Homo sapi
28	40.6	22.9	346259	1 BX640435	BX640435 Bordetell
29	40.6	22.9	346287	1 BX640450	BX640450 Bordetell
30	40.6	22.9	349672	1 BX640419	BX640419 Bordetell
31	40.4	22.8	717	8 AY488136	AY488136 Zea mays
32	40.4	22.8	283100	1 SC039110	SC039110 Streptomy
33	40.2	22.7	6182	3 AF236019	AF236019 Giardia i
34	40.2	22.7	300029	1 AE016787	AE016787 Pseudomon
35	40	22.6	285	6 AR009714	AR009714 Sequence
36	40	22.6	285	6 AR136764	AR136764 Sequence
37	40	22.6	678	12 AF506027	AF506027 Synthetic
38	40	22.6	299050	1 SC039104	SC039104 Streptomy
39	39.8	22.5	4940	9 AF180682	AF180682 Homo sapi
40	39.8	22.5	5409	6 CQ723772	CQ723772 Sequence
41	39.8	22.5	5434	6 AX330730	AX330730 Sequence
42	39.8	22.5	5434	6 AX330939	AX330939 Sequence
43	39.8	22.5	5434	6 AX332235	AX332235 Sequence
44	39.8	22.5	5434	9 HSDNMTASE	X63692 H.sapiens m
45	39.6	22.4	300511	1 AE016775	AE016775 Pseudomon

## ALIGNMENTS

RESULT 1	AF151798	354 bp	mRNA	linear	INV 06-OCT-2004
LOCUS	Mesobuthus martensii	insect beta-neurotoxin (bt)	mRNA, complete		
DEFINITION	Mesobuthus martensii				
ACCESSION	AF151798				
VERSION	AF151798.2	GI:53828929			
KEYWORDS	Mesobuthus martensii (Buthus martensii)				
SOURCE	Mesobuthus martensii				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Mesobuthus.				
REFERENCE	1 (bases 1 to 354)				
AUTHORS	Zeng,X.-C., Li,W.-X. and Zu,S.-Y.				
TITLE	A novel cDNA sequence encoding the precursor of a new type of insect beta-neurotoxin, BmKBT from Chinese scorpion Buthus martensii Karsch				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 354)				
AUTHORS	Li,W.-X., Zeng,X.-C. and Zu,S.-Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-MAY-1999) Department of Virology and Molecular Biology, Virology Institution of Wuhan University, LuoJia Street, Wuhan, Hubei 430072, P.R. China				
REFERENCE	3 (bases 1 to 354)				
AUTHORS	Li,W.-X., Zeng,X.-C. and Zu,S.-Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-OCT-2004) Department of Virology and Molecular Biology, Virology Institution of Wuhan University, LuoJia Street, Wuhan, Hubei 430072, P.R. China				
REMARK	Sequence update by submitter				
COMMENT	On Oct 6, 2004 this sequence version replaced gi.30844240.				
FEATURES	Location/Qualifiers				
source	1..354				
	/organism="Mesobuthus martensii"				
	/mol_type="mRNA"				
	/db_xref="taxon:34649"				
	/tissue_type="venom gland"				
gene	1..354				
	/notes="authority: Buthus martensii Karsch"				
CDS	12..254				

```

/gene="bt"
/note="BmKBT2"
/codon_start=1
/product="insect beta-neurotoxin"
/protein_id="AAP41418.2"
/db_xref="GI:53828930"
/translation="MMKFLFGFMFVILFSLMGSIIRGDDDPGNYPTNAYGNKYCTILG
ENEYCRKICKLHGVTYGYCYNRCWCEKLEKDVTI"

ORIGIN
Query Match          25.9%; Score 45.8; DB 3; Length 354;
Best Local Similarity 53.7%; Pred. No. 4.1;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1 GCCACGTCGCCAGGAACTACCCCTCGACAGCTCGGACACACACTACCTGTGGCAGCC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 78 GAGCAGATCGAGAAATTAATCGACAAATGCTTACGGTAACAATATTTATGTACGATT 137
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CTGGCGACAACCCGGACTGCATCAAGATCTGCCAGAACGACGCGTCGACTACGGCTAC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 138 TTGGGAGAAATGAATATTGTAGGAAAATATGTAATAATGTCATGAGTTACTTATGTTAT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TGTACGGCTTCAGTGTGGTGGCAGTTCTCTGAAGGACGACGACGTCAGGTTGTA 177
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 198 TGTACAAATTCGAGATCTGGTGCAGAAAATTTGGAAGATAAAGACGTCAACATTGA 254
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
AY282464
LOCUS AY282464 469 bp mRNA linear INV 01-MAY-2004
DEFINITION Mesobuthus martensii toxin KBT precursor, mRNA, complete cds.
ACCESSION AY282464
VERSION AY282464.1 GI:33590392
KEYWORDS
SOURCE Mesobuthus martensii (Buthus martensii)
ORGANISM Mesobuthus martensii
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
Buthida; Buthoidea; Buthidae; Mesobuthus.
REFERENCE 1 (bases 1 to 469)
AUTHORS Jiang,D., Cao,Z. and Li,W.
TITLE Cloning and characterizing of a new scorpion toxin from Buthus
martensii Karsch
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 469)
AUTHORS Jiang,D., Cao,Z. and Li,W.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2003) Biotechnology, College of Life Sciences,
Luojia Mountain, Wuhan, Hubei 430072, China
FEATURES
source
1..469
/organism="Mesobuthus martensii"
/mol_type="mRNA"
/db_xref="taxon:34649"
/tissue_type="venom"
/country="China"
/notes="authority: Mesobuthus martensii Karsch"
38..280
/codon_start=1
/product="toxin KBT precursor"
/protein_id="AA022733.1"
/db_xref="GI:33590393"
/translation="MMKFLFGFMFVILFSLMGSIIRGDDDPGNYPTNAYGNKYCTILG
ENEYCRKICKLHGVTYGYCYNRCWCEKLEKDVTI"
38..103
sig_peptide
mat_peptide
104..277
/product="toxin KBT"
329..334
polyA_signal

ORIGIN
Query Match          25.9%; Score 45.8; DB 3; Length 469;
Best Local Similarity 53.7%; Pred. No. 3.9;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```

```

/organism="Streptomyces avermitilis MA-4680"
/mol_type="genomic DNA"
/db_xref="taxon:227882"
/notes="This strain is also named as strain: ATCC 31267, NCIMB 12804 or NRRL 8165.-synonym: Streptomyces avermectinius"
106..1536
CDS
/feature="SAV5443"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC73155.1"
/db_xref="GI:29609104"
/translation="MTKDVIALTKMPDPLSLVLAGLSGGFDKLVGAEAGVAVQLCD
EGRPLVSVAPLLVQVGEARQLLGATEPEYVWMTAEARATGKVEAEQLAGTFAAR
LASLVGTAWPPQOAGSLAVVAKDGVRAAPTAQAQPAVDVLTDKVAVIQDRPVAM
TAWSLDALRAANAELGLQIVTPAGTSLSPAVRNSLPGWSPRVVQDEKGYDGLSG
AVLQWNGLFATVESPDATQDDHTPVADAFKKGLAEEGDTGSRQLAISRTIHPADD
RLVLGALSVWGRGITGEPGAGWTEPEANLPWSLRQLTDVAERSAPTFLVWVGSP
ERPLGATVRSRTTGGVEEDVTLAFGGYGGEEPPVDALPKVAETLATRHDLQSLVQI
RKARDLSVPPREGPGVPLAFALGAEVRAVMPQDRARRTPLAQPPVQLGPKTRPALY
YPLPGDPSDLSGWQDFRLMRHLKGA"
complement(1655..2782)
/feature="SAV5444"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC73156.1"
/db_xref="GI:29609105"
/translation="MPSFRAPQNSSSQISQTSRGASQAQSPESKPVKQRDAFEDNAK
YLAIVFVAMGHAEPLVDHRAAEALYVVTTHMPAFILISGYFSKSPDMRPDLRK
LVGTGAVPYLVFAYSLFKRWADNDQGHPISLDDPMLTFLVAFVWRLTLPWKV
VRVFPALAIAYLASVSPDIGDDLDLQVLQFLPFVVGFLFMKPEHQLLRREVRV
LSVPFALAIAYWAGPRMNSAWFYHRDQSAQELGAPWAGVWMTLALFGCSLVVTS
FPAVPRKMWFTALGAGTLGYLLHGFIAGKSRFDWDYDAWLTQPYGEVIATVPA
TLIVLCTPVQVIFRPMEPKMEWAFKRDATLARERAKA"
3488..3682
/feature="SAV5445"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC73157.1"
/db_xref="GI:29609106"
/translation="MAISLSVLLALLAILVLVLRGSSIKAGPAIVAVLFGFFLASTGM
APGINRPLDSIAETINSISF"
complement(4060..4133)
/feature="trn39"
/gene="trn39"
/complement(4060..4133)
/feature="trn39"
/gene="trn39"
/product="trnA-Gly"
/notes="codon recognized: GGA
trnA 39"
/anticodon=(pos:4099..4101,aa:Gly)
4287..4363
/feature="trn22"
4287..4363
/feature="trn22"
/feature="trn22"
/product="trnA-Pro"
/notes="codon recognized: CCA
trnA 22"
/anticodon=(pos:4321..4323,aa:Pro)
4549..5940
/feature="tig"
4549..5940
/feature="tig"
/feature="tig"
/feature="SAV5446"
/notes="SAV5446"
/codon_start=1
/transl_table=11
/product="putative cell division trigger factor"
/protein_id="BAC73158.1"
/db_xref="GI:29609107"
/translation="MKSAVETLNPTRVRLTVEVPFEELKDSLDAAAYKKINQVTVKGF

```

```

RKGIIPARVIDORFGRGAVLEAEVNDALPKFYTEAVNEAELNPLGOPEVDITELKQGE
TLNFTAEDVRPTIEIPDVGSGIEVEVDAVEVSDVEKAVTELARERAPSPVERAAE
KGDVVTLDLEAKVDGVEDGVAEGSYITIGSGELLDGIDDVAKGLAGAGETTTSEL
KGSAGAEAEVTVKVSVAARELPDLDDFAQLASEFDTLEELKADSRKLENNKQY
DQATQAEVRLEKLELAEPVPEKLEEDINTRKHNLEHQLQCMGLDLEKYLEIQG
KTVEFDAETKEAAVKGIKTOFVLDLVNKEKLVNQEBUTEHLMRRAASGSGSPDOF
QAAVRSQGVPMLAGVEARGKALAVVVEAATVKDTNGEIVDLDDDEDETESTPETTEA
EAAEESTEDKPEA"
7269..7928
/feature="clpP1"
/feature="clpP1"
/feature="clpP1"
/feature="SAV5447"
/codon_start=1
/transl_table=11
/product="putative ATP-dependent Clp protease proteolytic
subunit 1"
/protein_id="BAC73159.1"
/db_xref="GI:29609108"
/translation="MRPGAVVRAGGYVTNLMPTAAGDPGGGLGQVQVNRLLGERI
IFLQAVDDDIANKITIAQLLLSLDPKDIYLYINGPGGSITAGMAIYDTMQYIKNDV
VTIAMGMAAMQGFLLSAGTPGKRFPALPNAELIHOPAGLAGSASDIKTHAERLLHT
KKMAELTAFHTQTVEQVTRDSRDRWDFPVEAKEYGLIDDIIMPTAAGMPGGGGTGA
A"
8039..8719
/feature="clpP2"
/feature="clpP2"
/feature="SAV5448"
/codon_start=1
/transl_table=11
/product="putative ATP-dependent Clp protease proteolytic
subunit 2"
/protein_id="BAC73160.1"
/db_xref="GI:29609109"
/translation="MNDYPSGLHARTQAEYTCARAESRVVIRFVERTSQGVREYDP
YAKLPERVIFLQVQDDASANDVMAQLLCLEMSDDRDSDISVINSPGGSFTALTAY
DTMQVKPDITQVCMGOASAAAILLAAGTPGKRMPALPNARVLIHOPYSTGRGQVSD
LETAANEILMRSSQLEDMLAKHSTTPIEKIREDIRDKILTAEADALAYGLIDQIISTR
KMNSAVA"
8880..10166
/feature="clpX2"
/feature="clpX2"
/feature="SAV5449"
/codon_start=1
/transl_table=11
/product="putative ATP-dependent Clp Protease ATP binding
subunit"
/protein_id="BAC73161.1"
/db_xref="GI:29609110"
/translation="MARIQDGLLKCFCGSKQKVKKLIAGPGVVICDECIDLQNE
IIEELAETSEVRWEELPKPREIYEFLEGVVGEEAAKALSAVNVNHYKRVQAGENG
GQGSREDAIELAKSNILLGPTSGKTLAQTLARMNVPAIADATALTAEAGVSGD
VENILKLIQAADYDVKKAEITGIYIDEDKVAKSENPSITRDVSGEGVQALLKIL
EGTASVPPGGGRKHQBFIQIDTNNVLFVGGAGAGLEKLIERSAGAKGIGFAT
RSKRELESQKQFEDVMPDLVKFGMIPEFIRGLPVITSVHNLREALLQLVLEPRNAL
VKQORLFDGLDVELDFEREALAIADAILROTGARGLRAIMEEVLMSVMYEVPSRK
DVARVITADVHNSNVNPTLIPRDARGRGPGEQKTA"
complement(10277..11245)
/feature="SAV5450"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC73162.1"
/db_xref="GI:29609111"
/translation="MSFNQPGPYGGPQQPGPYGQPGPYQPPQAPQPGYGYGQQAQ
QPGYGYGQQAQPPQPGPYGQPGYQPPQPGPYGQPPQPPAPGPGGKKTKGLII
GAVAVVAAVAGAYFVIGSGGSLADGGAHKLTTPTVLSEYKRVGKGGSSSNSTV

```

Query Match 24.7%; Score 43.8; DB 1; Length 299925;  
 Best Local Similarity 60.5%; Pred. No. 5.5;  
 Matches 72; Conservative 0; Mismatches 47; Indels 0; Gaps 0;





```
ORIGIN
/db_xref="taxon:32630"
/note="non-aggregating mutant"

Query Match      23.5%; Score 41.6; DB 6; Length 675;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGAGCTGCCAGGGAACCTACCCGCTGGACACTCGGACAAACACCTACTCTGTGCGACCCC 61
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CCAAGGTGTACGTGAAGCACCACCCCGGACATCCCGACTACAAGAAGCTGTCTTCCCG 265

Qy 62 TGGCGACAACCCGAGCTGCATCAAGATCTCCAGAGCAGCGGTGACTACGGTACT 121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 AGGCTTCAAGTGGAGCGGTGATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325

Qy 122 GCTACGGCTTCCAGTGTGGTGGCGAGTTCTTGAAGGACGAGAAGCTCA 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AGGACTCTCTCCGAGGACGGCTTTCATCTACAAGGTGAAGTTCA 373

RESULT 7
AX370404
LOCUS AX370404 678 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 1 from Patent WO0196373.
ACCESSION AX370404
VERSION AX370404.1 GI:18857490
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Fradkov,A.F. and Tersikh,A.
TITLE Fluorescent timer proteins and methods for their use
JOURNAL Patent: WO 0196373-A 1 20-DEC-2001;
Clontech Laboratories Inc. (US)
FEATURES
source
1..678
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant of sequence from Discosoma sp."

ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGAGCTGCCAGGGAACCTACCCGCTGGACACTCGGACAAACACCTACTCTGTGCGACCCC 61
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CCAAGGTGTACGTGAAGCACCACCCCGGACATCCCGACTACAAGAAGCTGTCTTCCCG 265

Qy 62 TGGCGACAACCCGAGCTGCATCAAGATCTCCAGAGCAGCGGTGACTACGGTACT 121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 AGGCTTCAAGTGGAGCGGTGATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325

Qy 122 GCTACGGCTTCCAGTGTGGTGGCGAGTTCTTGAAGGACGAGAAGCTCA 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AGGACTCTCTCCGAGGACGGCTTTCATCTACAAGGTGAAGTTCA 373

RESULT 8
AX370406
LOCUS AX370406 678 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 3 from Patent WO0196373.
ACCESSION AX370406
VERSION AX370406.1 GI:18857491
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Fradkov,A.F. and Tersikh,A.

TITLE Fluorescent timer proteins and methods for their use
JOURNAL Patent: WO 0196373-A 5 20-DEC-2001;
Clontech Laboratories Inc. (US)
FEATURES
source
1..678
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant of sequence from Discosoma sp."

ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGAGCTGCCAGGGAACCTACCCGCTGGACACTCGGACAAACACCTACTCTGTGCGACCCC 61
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CCAAGGTGTACGTGAAGCACCACCCCGGACATCCCGACTACAAGAAGCTGTCTTCCCG 265

Qy 62 TGGCGACAACCCGAGCTGCATCAAGATCTCCAGAGCAGCGGTGACTACGGTACT 121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 AGGCTTCAAGTGGAGCGGTGATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325

Qy 122 GCTACGGCTTCCAGTGTGGTGGCGAGTTCTTGAAGGACGAGAAGCTCA 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AGGACTCTCTCCGAGGACGGCTTTCATCTACAAGGTGAAGTTCA 373

RESULT 9
AX370408
LOCUS AX370408 678 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 5 from Patent WO0196373.
ACCESSION AX370408
VERSION AX370408.1 GI:18857492
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Fradkov,A.F. and Tersikh,A.
TITLE Fluorescent timer proteins and methods for their use
JOURNAL Patent: WO 0196373-A 5 20-DEC-2001;
Clontech Laboratories Inc. (US)
FEATURES
source
1..678
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant of sequence from Discosoma sp."

ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGAGCTGCCAGGGAACCTACCCGCTGGACACTCGGACAAACACCTACTCTGTGCGACCCC 61
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CCAAGGTGTACGTGAAGCACCACCCCGGACATCCCGACTACAAGAAGCTGTCTTCCCG 265

Qy 62 TGGCGACAACCCGAGCTGCATCAAGATCTCCAGAGCAGCGGTGACTACGGTACT 121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 AGGCTTCAAGTGGAGCGGTGATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325

Qy 122 GCTACGGCTTCCAGTGTGGTGGCGAGTTCTTGAAGGACGAGAAGCTCA 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 AGGACTCTCTCCGAGGACGGCTTTCATCTACAAGGTGAAGTTCA 373

RESULT 10
AX824725
LOCUS AX824725 678 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 7 from Patent WO02068459.
ACCESSION AX824725
```

```

VERSION      AX824725.1  GI:39750591
KEYWORDS     Discosoma sp.
SOURCE       Discosoma sp.
ORGANISM     Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
              Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE    1
AUTHORS      Non aggregating fluorescent proteins and methods for using the same
TITLE        Patent: WO 0208459-A 7 06-SEP-2002;
JOURNAL      Location/Qualifiers
FEATURES     source
              1..678
                /organism="Discosoma sp."
                /mol_type="unassigned DNA"
                /db_xref="taxon:86600"
ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACACCTACCTGTGGCACCCC 61
Db 206 CCAAGGTGTACGTGAAGCAACCCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 265
Qy 62 TGGGCGCAACCCCGGACTGCATCAAGATCTGCCAGAAAGCAACGGCGTGCAGTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGGTGATGAATCTCGNGAGCGGGCGGTGGTACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
Db 326 AGGACTCCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCA 373

RESULT 11
AX824732
LOCUS        AX824732              678 bp DNA linear PAT 11-DEC-2003
DEFINITION   Sequence 14 from Patent WO02068459.
ACCESSION   AX824732
VERSION     AX824732.1  GI:39750595
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     other sequences; artificial sequences.
REFERENCE    1
AUTHORS      Non aggregating fluorescent proteins and methods for using the same
TITLE        Patent: WO 0208459-A 14 06-SEP-2002;
JOURNAL      Location/Qualifiers
FEATURES     source
              1..678
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="non-aggregating mutant"
ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 678;
Best Local Similarity 53.0%; Pred. No. 32;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACACCTACCTGTGGCACCCC 61
Db 206 CCAAGGTGTACGTGAAGCAACCCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 265
Qy 62 TGGGCGCAACCCCGGACTGCATCAAGATCTGCCAGAAAGCAACGGCGTGCAGTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGGTGATGAATCTTCGAGGACGCGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
Db 326 AGGACTCCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCA 373

```

```

RESULT 12
AR527331
LOCUS        AR527331              723 bp DNA linear PAT 08-OCT-2004
DEFINITION   Sequence 1 from patent US 6723537.
ACCESSION   AR527331
VERSION     AR527331.1  GI:53914309
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 723)
AUTHORS      Peeille,B.
TITLE        Directed evolution of protein in mammalian cells
JOURNAL      Patent: US 6723537-A 1 20-APR-2004;
FEATURES     Location/Qualifiers
              source
                1..723
                  /organism="unknown"
                  /mol_type="genomic DNA"
ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 723;
Best Local Similarity 53.0%; Pred. No. 31;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACACCTACCTGTGTGGCACCCC 61
Db 209 CCAAGGTGTACGTGAAGCAACCCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 268
Qy 62 TGGGCGCAACCCCGGACTGCATCAAGATCTGCCAGAAAGCAACGGCGTGCAGTACGGCTACT 121
Db 269 AGGCTTCAAGTGGAGCGGTGATGAATCTTCGAGGACGCGGCGGTGGTACCGTGACCC 328
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
Db 329 AGGACTCCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCA 376

RESULT 13
AX666133
LOCUS        AX666133              1050 bp DNA linear PAT 26-MAR-2003
DEFINITION   Sequence 7 from Patent WO02060941.
ACCESSION   AX666133
VERSION     AX666133.1  GI:29290961
KEYWORDS     .
SOURCE       unidentified
ORGANISM     unclassified.
REFERENCE    1
AUTHORS      Zhao,M., Xu,M., Jiang,P. and Yang,M.
TITLE        Fluorescent proteins
JOURNAL      Patent: WO 02060941-A 7 08-AUG-2002;
              Anticancer, Inc. (US)
FEATURES     Location/Qualifiers
              source
                1..1050
                  /organism="unidentified"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32644"
                  /note="Coral"
ORIGIN
Query Match      23.5%; Score 41.6; DB 6; Length 1050;
Best Local Similarity 53.0%; Pred. No. 30;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACACCTACCTGTGTGGCACCCC 61
Db 494 CCAAGGTGTACGTGAAGCAACCCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 553
Qy 62 TGGGCGCAACCCCGGACTGCATCAAGATCTGCCAGAAAGCAACGGCGTGCAGTACGGCTACT 121
Db 554 AGGCTTCAAGTGGAGCGGTGATGAATCTTCGAGGACGCGGCGGTGGGACCGTGACCC 613
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169

```

```
Db 614 AGGACTCTCCCTCGAGACGGCTGCTTCTACTACAGGTGAAGTTCA 661
|||||
RESULT 14
CQ882115 2721 bp DNA linear PAT 11-OCT-2004
LOCUS Sequence 1 from Patent WO2004083445.
DEFINITION CQ882115
ACCESSION CQ882115.1 GI:54034825
VERSION CQ882115.1
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1
REFERENCE 1
AUTHORS Chavanc'y.G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,
Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.
TITLE Nucleic acid controlling the expression of a useful polypeptide in
the posterior silk glands of a lepidoptera and application thereof
JOURNAL Patent: WO 2004083445-A 1 30-SEP-2004;
Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE
CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche
Agronomique (INRA) (FR)
FEATURES
source
Location/Qualifiers
1..2721
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Description de la s quence artificielle : s quence
de fusion"

ORIGIN
Query Match 23.5%; Score 41.6; DB 6; Length 2721;
Best Local Similarity 53.0%; Pred. No. 27;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTCGACGGGAACCTACCCGCTGGACAGCTCGGACACACCTACTGTGCGCACCCC 61
Db 2300 CCAAGGTGTACGTGAAGCAGCAGCCCGCCGACATCCCGACTACAAGAGCTGTCTTCGCCG 2308
Qy 62 TGGCGCACACCCCGACTGCATCAAGATCTGCCAGAGCAGCGGTGCTACTACGGCTACT 121
Db 2360 AGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAGGACGGCGGTGGTACCGTGACCC 2419
Qy 122 GCTACGCGTTCACAGTGTGGTGGAGTTCCTCTGAAGGACGAGAACGTCA 169
Db 2420 AGGACTCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCA 2467

Search completed: July 30, 2005, 19:11:14
Job time : 1269.87 secs

Db 614 AGGACTCTCCCTCGAGACGGCTGCTTCTACTACAGGTGAAGTTCA 661
|||||
RESULT 15
CQ882117 2772 bp DNA linear PAT 11-OCT-2004
LOCUS Sequence 3 from Patent WO2004083445.
DEFINITION CQ882117
ACCESSION CQ882117
VERSION CQ882117.1 GI:54034827
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1
REFERENCE 1
AUTHORS Chavanc'y.G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,
Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.
TITLE Nucleic acid controlling the expression of a useful polypeptide in
the posterior silk glands of a lepidoptera and application thereof
JOURNAL Patent: WO 2004083445-A 3 30-SEP-2004;
Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE
CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche
Agronomique (INRA) (FR)
FEATURES
source
Location/Qualifiers
1..2772
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Description de la s quence artificielle : s quence
de fusion"

ORIGIN
Query Match 23.5%; Score 41.6; DB 6; Length 2721;
Best Local Similarity 53.0%; Pred. No. 27;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 2 CCGACGTCGACGGGAACCTACCCGCTGGACAGCTCGGACACACCTACTGTGCGCACCCC 61
Db 2249 CCAAGGTGTACGTGAAGCAGCAGCCCGCCGACATCCCGACTACAAGAGCTGTCTTCGCCG 2308
Qy 62 TGGCGCACACCCCGACTGCATCAAGATCTGCCAGAGCAGCGGTGCTACTACGGCTACT 121
Db 2309 AGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAGGACGGCGGTGGTACCGTGACCC 2368
Qy 122 GCTACGCGTTCACAGTGTGGTGGAGTTCCTCTGAAGGACGAGAACGTCA 169
Db 2369 AGGACTCTCCCTCGAGACGGCTGCTTCTACTACAGGTGAAGTTCA 2416

RESULT 15
CQ882117 2772 bp DNA linear PAT 11-OCT-2004
LOCUS Sequence 3 from Patent WO2004083445.
DEFINITION CQ882117
ACCESSION CQ882117
VERSION CQ882117.1 GI:54034827
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1
REFERENCE 1
AUTHORS Chavanc'y.G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,
Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.
TITLE Nucleic acid controlling the expression of a useful polypeptide in
the posterior silk glands of a lepidoptera and application thereof
JOURNAL Patent: WO 2004083445-A 3 30-SEP-2004;
Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE
CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche
Agronomique (INRA) (FR)
FEATURES
source
Location/Qualifiers
1..2772
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Description de la s quence artificielle : s quence
de fusion"

ORIGIN
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 18:20:14 ; Search time 389.677 Seconds  
(without alignments)  
2973.108 Million cell updates/sec

Title: US-10-617-978-14\_COPY\_62\_240

Perfect score: 179

Sequence: 1 cggctgcgctccgggaac.....atgagacgtgaaggtctga 179

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA: \*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description
1	38	21.2	189	21 US-10-721-793-79
2	38	21.2	316	21 US-10-721-793-77
3	36.8	20.6	195	21 US-10-721-793-139
4	36.8	20.6	195	21 US-10-721-793-143
5	36.8	20.6	195	21 US-10-721-793-147
6	36.8	20.6	195	21 US-10-721-793-167
7	36.8	20.6	195	21 US-10-721-793-171

8	36.8	20.6	198	21	US-10-721-793-155	Sequence 155, App
9	36.8	20.6	198	21	US-10-721-793-159	Sequence 159, App
10	36.8	20.6	198	21	US-10-721-793-163	Sequence 163, App
11	36.8	20.6	319	21	US-10-721-793-153	Sequence 153, App
12	36.8	20.6	323	21	US-10-721-793-137	Sequence 137, App
13	36.8	20.6	323	21	US-10-721-793-141	Sequence 141, App
14	36.8	20.6	323	21	US-10-721-793-145	Sequence 145, App
15	36.8	20.6	323	21	US-10-721-793-157	Sequence 157, App
16	36.8	20.6	323	21	US-10-721-793-161	Sequence 161, App
17	36.8	20.6	323	21	US-10-721-793-165	Sequence 165, App
18	36.8	20.6	323	21	US-10-721-793-169	Sequence 169, App
19	36.4	20.3	219	21	US-10-721-793-15	Sequence 15, App
20	35.2	19.7	273	21	US-10-721-793-13	Sequence 13, Appl
21	35.2	19.7	198	21	US-10-721-793-175	Sequence 175, App
22	35.2	19.7	323	21	US-10-721-793-173	Sequence 173, App
23	34.8	19.4	189	21	US-10-721-793-39	Sequence 39, Appl
24	34.8	19.4	189	21	US-10-721-793-51	Sequence 51, Appl
25	34.8	19.4	189	21	US-10-721-793-111	Sequence 111, App
26	34.8	19.4	252	21	US-10-721-793-49	Sequence 49, Appl
27	34.8	19.4	254	21	US-10-721-793-37	Sequence 37, Appl
28	34.8	19.4	311	21	US-10-721-793-109	Sequence 109, App
29	34.6	19.3	195	21	US-10-721-793-3	Sequence 3, Appli
30	34.6	19.3	195	21	US-10-721-793-7	Sequence 7, Appli
31	34.6	19.3	195	21	US-10-721-793-27	Sequence 27, Appl
32	34.6	19.3	195	21	US-10-721-793-35	Sequence 35, Appl
33	34.6	19.3	261	21	US-10-721-793-25	Sequence 25, Appl
34	34.6	19.3	261	21	US-10-721-793-33	Sequence 33, Appl
35	34.6	19.3	299	21	US-10-721-793-1	Sequence 1, Appli
36	34.6	19.3	317	21	US-10-721-793-5	Sequence 5, Appli
37	33.6	18.8	195	21	US-10-721-793-23	Sequence 23, Appl
38	33.6	18.8	195	21	US-10-721-793-31	Sequence 31, Appl
39	33.6	18.8	204	21	US-10-721-793-19	Sequence 19, Appl
40	33.6	18.8	261	21	US-10-721-793-17	Sequence 17, Appl
41	33.6	18.8	261	21	US-10-721-793-29	Sequence 29, Appl
42	33.6	18.8	314	21	US-10-721-793-21	Sequence 21, Appl
43	33.4	18.7	189	21	US-10-721-793-187	Sequence 187, App
44	33.4	18.7	314	21	US-10-721-793-185	Sequence 185, App
45	32.2	18.0	192	21	US-10-721-793-103	Sequence 103, App

#### ALIGNMENTS

#### RESULT 1

```
US-10-721-793-79
; Sequence 79, Application US/10721793
; Publication No. US2005006531A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immungens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721.793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 79
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Centruroides limpidus limpidus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; US-10-721-793-79
```



```
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
; US-10-721-793-143

Query Match      20.6%; Score 36.8; DB 21; Length 195;
Best Local Similarity 73.4%; Pred. No. 0.0091;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 157 AGGA 160
   |||||
Db 155 CCGA 158

RESULT 5
US-10-721-793-147
; Sequence 147, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721.793
; CURRENT FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
; US-10-721-793-147

Query Match      20.6%; Score 36.8; DB 21; Length 195;
Best Local Similarity 73.4%; Pred. No. 0.0091;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 157 AGGA 160
   |||||
Db 155 CCGA 158
```

```
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
; US-10-721-793-147

Query Match      20.6%; Score 36.8; DB 21; Length 195;
Best Local Similarity 73.4%; Pred. No. 0.0091;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 157 AGGA 160
   |||||
Db 155 CCGA 158

RESULT 6
US-10-721-793-167
; Sequence 167, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721.793
; CURRENT FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Centruroides sculpuratus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(195)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpuratus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(195)
; US-10-721-793-167

Query Match      20.6%; Score 36.8; DB 21; Length 195;
Best Local Similarity 73.4%; Pred. No. 0.0091;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGGAATTTCTGA 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 157 AGGA 160
   |||||
Db 155 CCGA 158

RESULT 7
US-10-721-793-171
```

```

Query Match          20.6%; Score 36.8; DB 21; Length 198;
Best Local Similarity 73.4%; Pred. NO. 0.0092;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAACAACGGTGTGGATTACGGGTATTGTACGCCCTTCCAATGCTGGTGTGAATTTCTGA 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 AGAACCAACGAGGTAGTTACGGGTATTGTACGCTTTCGCATGCTGGTGCGAAGGTTTC 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 157 AGGA 160
    |||
Db 155 CCGA 158
    |||

RESULT 9
US-10-721-793-159
; Sequence 159, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:

```

```

/ NUMBER OF SEQ ID NOS: 231
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 159
/ LENGTH: 198
/ TYPE: DNA
/ ORGANISM: Centruroides sculpturatus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(198)
/ OTHER INFORMATION: Product= Sodium-channel modifier toxin
/ PUBLICATION INFORMATION:
/ AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
/ TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
/ TITLE: that recognize Na+-channels
/ JOURNAL: Toxicon
/ VOLUME: 39
/ ISSUE: 12
/ PAGES: 1893-1898
/ DATE: 2001-12-01
/ DATABASE ENTRY DATE:
/ RELEVANT RESIDUES: (1)..(198)
/ US-10-721-793-159

```



```
Query Match      20.6%; Score 36.8; DB 21; Length 198;
Best Local Similarity 73.4%; Pred. No. 0.0092;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAACACGGTGTGGATTACGGGTATTGCTACGGCTTCCCAATGCTGGTGAATTTCTGA 156
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGGCTTTCGTCATGCTGGTGCAGAGGTTTC 154
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 157 AGGA 160
    ||
Db 155 CCGA 158

RESULT 10
US-10-721-793-163
; Sequence 163, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 163
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(198)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(198)
US-10-721-793-163

Query Match      20.6%; Score 36.8; DB 21; Length 198;
Best Local Similarity 73.4%; Pred. No. 0.0092;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAACACGGTGTGGATTACGGGTATTGCTACGGCTTCCCAATGCTGGTGAATTTCTGA 156
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGGCTTTCGTCATGCTGGTGCAGAGGTTTC 154
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 157 AGGA 160
    ||
Db 155 CCGA 158

RESULT 11
US-10-721-793-153
; Sequence 153, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 153
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(261)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (1)..(261)
US-10-721-793-153

Query Match      20.6%; Score 36.8; DB 21; Length 319;
Best Local Similarity 73.4%; Pred. No. 0.011;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAACACGGTGTGGATTACGGGTATTGCTACGGCTTCCCAATGCTGGTGAATTTCTGA 156
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGGCTTTCGTCATGCTGGTGCAGAGGTTTC 211
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 157 AGGA 160
    ||
Db 212 CCGA 215

RESULT 12
US-10-721-793-137
; Sequence 137, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
```

```
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 137
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-137

Query Match 20.6%; Score 36.8; DB 21; Length 323;
Best Local Similarity 73.4%; Pred. No. 0.011;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAACACGGTGTGGATTACGGGTATTGCTACGCCITCCCAATGCTGTGTAATTTCTGA 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 AGAACCAAGGAGGAGTGTACGGGTATTGCTACGCCITTCGCATGCTGTGCGAAGGTTTC 215
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 157 AGGA 160
||
Db 216 CCGA 219

RESULT 13
US-10-721-793-141
; Sequence 141, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
```

```
; TITLE OF INVENTION: Venom of Scorpions of the Genus Centruroides
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Centruroides sculpturatus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(265)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Cys is amidated, and the last Gly
; OTHER INFORMATION: and the last 2 basic aminoacids are cut
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (62)..()
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (269)..(323)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Possani L.D.
; TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
; TITLE: that recognize Na+-channels
; JOURNAL: Toxicon
; VOLUME: 39
; ISSUE: 12
; PAGES: 1893-1898
; DATE: 2001-12-01
; DATABASE ENTRY DATE:
; RELEVANT RESIDUES: (5)..(265)
US-10-721-793-141

Query Match 20.6%; Score 36.8; DB 21; Length 323;
Best Local Similarity 73.4%; Pred. No. 0.011;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAACACGGTGTGGATTACGGGTATTGCTACGCCITCCCAATGCTGTGTAATTTCTGA 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 AGAACCAAGGAGGAGTGTACGGGTATTGCTACGCCITTCGCATGCTGTGCGAAGGTTTC 215
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 157 AGGA 160
||
Db 216 CCGA 219

RESULT 14
US-10-721-793-145
; Sequence 145, Application US/10721793
; Publication No. US20050065331A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
```

```

: CURRENT APPLICATION NUMBER: US/10/721,793
: CURRENT FILING DATE: 2003-11-26
: PRIOR APPLICATION NUMBER: US 60/430,067
: PRIOR FILING DATE: 2002-12-02
: NUMBER OF SEQ ID NOS: 294
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 157
: LENGTH: 323
: TYPE: DNA
: ORGANISM: Centruroides sculpturatus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (5)..(265)
: OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
: OTHER INFORMATION: In the mature peptide, the last 2 basic aminoacids are cut
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: (269)..(323)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: (1)..(4)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: (5)..(61)
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: (62)..()
: OTHER INFORMATION: Product= Sodium-channel modifier toxin
: PUBLICATION INFORMATION:
: AUTHORS: Corona, M., Valdez-Cruz, N.A., Merino, E., Zurita, M. & Poesani L.D.
: TITLE: Genes and peptides from the scorpion Centruroides sculpturatus Ewing,
: TITLE: that recognize Nat-channels
: JOURNAL: Toxicon
: VOLUME: 39
: ISSUE: 12
: PAGES: 1893-1898
: DATE: 2001-12-01
: DATABASE ENTRY DATE:
: RELEVANT RESIDUES: (5)..(265)
: US-10-721-793-157

Query Match 20.6%; Score 36.8; DB 21; Length 323;
Best Local Similarity 73.4%; Pred. No. 0.011;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 97 AGAAACACGGTGGGATTACGGGTATTGCTACGCCCTTCCAAATGCTGTGTAATTCTGA 156
DB 156 AGAACCAAGAGTAGTACGGGTATTGCTACGCTTTCGCATGCTGTGCGAAGTTTGC 215
QY 157 AGGA 160
DB 216 CCGA 219

Search completed: July 30, 2005, 20:32:11
Job time : 390.677 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 17:12:09 ; Search time 85.4775 Seconds  
(without alignments)  
3426.556 Million cell updates/sec

Title: US-10-617-978-14\_COPY\_62\_240  
Perfect score: 179  
Sequence: 1 cggctgacgtcccggaac.....atgagaacgtgaaggtctga 179

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgm2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgm2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgm2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgm2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgm2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42.2	23.6	270	4	US-09-599-632-34
2	40.8	22.8	270	4	US-09-599-632-32
3	40.8	22.8	270	4	US-09-599-632-36
4	38	21.2	270	4	US-09-599-632-30
C 5	32.6	18.2	336	3	US-08-931-858E-97
C 6	32.6	18.2	336	3	US-08-981-739-97
C 7	32.6	18.2	336	3	US-08-981-739-106
C 8	32.6	18.2	336	3	US-08-128-026-97
C 9	32.6	18.2	336	3	US-09-128-026-106
C 10	32.6	18.2	336	4	US-09-220-616-97
C 11	32.6	18.2	336	4	US-09-220-616-106
C 12	32.6	18.2	336	4	US-09-220-527-97
C 13	32.6	18.2	336	4	US-09-220-527-106
C 14	32.6	18.2	336	4	US-09-220-407-97
C 15	32.6	18.2	391	3	US-08-931-858E-107
C 16	32.6	18.2	391	3	US-08-981-739-107
C 17	32.6	18.2	391	3	US-09-128-026-107
C 18	32.6	18.2	391	4	US-09-220-616-107
C 19	32.6	18.2	391	4	US-09-220-527-107
C 20	32.6	18.2	391	4	US-09-220-407-107
21	30.2	16.9	3379	4	US-09-220-132-12
22	30.2	16.9	3445	4	US-09-976-594-323
C 23	30	16.8	42348	4	US-09-949-016-17157
C 24	29.8	16.6	726	4	US-09-702-705-1344
C 25	29.8	16.6	726	4	US-09-736-457-1344
C 26	29.8	16.6	726	4	US-09-614-124B-1344
C 27	29.8	16.6	726	4	US-09-671-325-1344

C 28	29.8	16.6	726	4	US-09-658-824-1344
C 29	29.2	16.3	40352	3	US-08-846-1110-15
C 30	29.2	16.3	40352	3	US-09-443-077-15
31	29	16.2	900	4	US-09-393-634-44
32	29	16.2	900	4	US-09-949-016-1071
33	29	16.2	900	4	US-09-949-016-2285
34	29	16.2	900	4	US-09-949-016-12813
35	29	16.2	4900	4	US-09-949-016-14027
36	28.8	16.1	301	3	US-09-053-021-3
37	28.8	16.1	345	3	US-09-053-021-8
38	28.2	15.8	601	4	US-09-949-016-52321
39	28.2	15.8	793	2	US-08-467-603-69
40	28.2	15.8	793	2	US-08-466-793-69
41	28.2	15.8	793	2	US-08-431-861A-69
42	28.2	15.8	793	4	US-09-374-671A-69
C 43	28.2	15.8	2456	1	US-07-882-711-1
C 44	28.2	15.8	2456	2	US-08-462-174-1
C 45	28.2	15.8	2486	4	US-09-949-016-4380

ALIGNMENTS

RESULT 1

US-09-599-632-34  
; Sequence 34, Application US/09599632  
; Patent No. 6768002  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1375 US NA  
; CURRENT APPLICATION NUMBER: US/09/599,632  
; CURRENT FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,410  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 34  
; LENGTH: 270  
; TYPE: DNA  
; ORGANISM: Hottentotta judiaca  
US-09-599-632-34

Query Match 23.6%; Score 42.2; DB 4; Length 270;  
Best Local Similarity 61.3%; Pred. No. 5.5e-06;  
Matches 68; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy	67	GAGATAATCCGGACTGCATTAAAGATCTGTGAGAACACGGTGTGGATTACGGGTATTGCT	126
Db	128	GTGATCATGATTATTGTGCGACATTTGTAAGATACATGAGAGTGAATTTATGGGTATTGTT	187

Qy	127	ACGCTTCCCATGCTGGTGTGAATTTCTGAGGATGAGACGTGACGTCT	177
Db	188	GGGTCACTCGTGTGTGTAATTTTGAAGAGAGACATCAATATTT	238

RESULT 2

US-09-599-632-32  
; Sequence 32, Application US/09599632  
; Patent No. 6768002  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1375 US NA  
; CURRENT APPLICATION NUMBER: US/09/599,632  
; CURRENT FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,410  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 32  
; LENGTH: 270  
; TYPE: DNA  
; ORGANISM: Hottentotta judiaca  
; US-09-599-632-32

Query Match 22.8%; Score 40.8; DB 4; Length 270;  
Best Local Similarity 61.1%; Pred. No. 1.9e-05;  
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATATCCGGACTGCAATTAAGATCTGTCAGAAACACGGTGTGGATTACGGGTATTGCTACG 129  
|||  
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGAGTGAATTATGGGTATTGTTGGG 190  
|||

Qy 130 CCTTCCAATCGTGTGGAATTTCTGAAGGATGAGACGTAAGGTCT 177  
|||  
Db 191 TCACCTCGTGTGGTGAATATTGTAAGAAGAAGACATCAATATTT 238  
|||

RESULT 3  
US-09-599-632-36  
; Sequence 36, Application US/09599632  
; Patent No. 6768002  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1375 US NA  
; CURRENT APPLICATION NUMBER: US/09/599,632  
; CURRENT FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,410  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 36  
; LENGTH: 270  
; TYPE: DNA  
; ORGANISM: Hottentotta judiaca  
; US-09-599-632-36

Query Match 22.8%; Score 40.8; DB 4; Length 270;  
Best Local Similarity 61.1%; Pred. No. 1.9e-05;  
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATATCCGGACTGCAATTAAGATCTGTCAGAAACACGGTGTGGATTACGGGTATTGCTACG 129  
|||  
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGAGTGAATTATGGGTATTGTTGGG 190  
|||

Qy 130 CCTTCCAATCGTGTGGAATTTCTGAAGGATGAGACGTAAGGTCT 177  
|||  
Db 191 TCACCTCGTGTGGTGAATATTGTAAGAAGAAGACATCAATATTT 238  
|||

RESULT 4  
US-09-599-632-30  
; Sequence 30, Application US/09599632  
; Patent No. 6768002  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1375 US NA  
; CURRENT APPLICATION NUMBER: US/09/599,632  
; CURRENT FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,410  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 30  
; LENGTH: 270

; TYPE: DNA  
; ORGANISM: Hottentotta judiaca  
; US-09-599-632-30

Query Match 21.2%; Score 38; DB 4; Length 270;  
Best Local Similarity 56.0%; Pred. No. 0.00022;  
Matches 93; Conservative 0; Mismatches 70; Indels 3; Gaps 1;

Qy 12 CCGGGAAACTACCACCTTGATTCTCCGACAATACCTACTGTGCGCCCTTTGGGAGAT 71  
|||  
Db 76 CAGGAAATACCAGATATCTGTTATGGTACGTCTTATGGATGCACAGCTTTTATCAT 135  
|||

Qy 72 AATCCGAGCTGCATTAAGATCTGTCAAGAACACGGTGTGGATTACGGGTATTGCTACGCC 131  
|||  
Db 136 AAT--TATTGTGGACATTTGTAAGTACATGAGTAAAGTATCGGTATTGTTGGGTC 192  
|||

Qy 132 TTCCAATGCTGGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177  
|||  
Db 193 ACCTCGTGTGGTGAATATTGTAAGAAAGAGACATCGATATTT 238  
|||

RESULT 5  
US-08-931-858E-97/c  
; Sequence 97, Application US/08931858E  
; Patent No. 6222022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931,858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; US-08-931-858E-97

Query Match 18.2%; Score 32.6; DB 3; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.03;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 25 CACTTGATTCTCCGACAATACCTACTGTGCGCCCTTTGGGAGATAATCCGACTGCA 84  
|||  
Db 285 CACCTGAGCCTTTCAAAGGCTGCTTGTGTTGATCTTGAGAGACCAGGCTGGCGCC 226  
|||

Qy 85 TTAGATCTGTCAACAACGGTGTGATTACGGGTATTGCTAGCGCTTCCAAATGCTGGT 144  
Db 225 TTACGCCACCACACCCAGCTCGGCTGAGAGCTGAGCGAGCTGCTGCCAATGGTGGT 166  
Qy 145 GTGAATTTCTGAAGG 159  
Db 165 GGTCAATCAAGGAAGG 151

## RESULT 6

US-08-981-739-97/c  
; Sequence 97, Application US/08981739  
; Patent No. 6232449  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; MILBRANDT, JEFFREY D.  
; KOTZBAUER, PAUL T.  
; LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981,739  
; FILING DATE: 31-Aug-1998  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/03461  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976163  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:  
US-08-981-739-97

Query Match 18.2%; Score 32.6; DB 3; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.03;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 25 CACTTGATTTCTCGACAATACCTGCTGCGCCCTTTGGGAGATAATCCGACTGCA 84  
Db 285 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGATTCTGAGAGACAGGCTGGCGGCC 226  
Qy 85 TTAGATCTGTCAACAACGGTGTGATTACGGGTATTGCTAGCGCTTCCAAATGCTGGT 144  
Db 225 TTACGCCACCACACCCAGCTCGGCTGAGAGCTGAGCGAGCTGCTGCCAATGGTGGT 166  
Qy 145 GTGAATTTCTGAAGG 159  
Db 165 GGTCAATCAAGGAAGG 151

## RESULT 7

US-08-981-739-106/c  
; Sequence 106, Application US/08981739  
; Patent No. 6232449  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; MILBRANDT, JEFFREY D.  
; KOTZBAUER, PAUL T.  
; LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981,739  
; FILING DATE: 31-Aug-1998  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/03461  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 976163  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 106:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 106:  
US-08-981-739-106

Query Match 18.2%; Score 32.6; DB 3; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.03;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 25 CACTTGATTTCTCGACAATACCTGCTGCGCCCTTTGGGAGATAATCCGACTGCA 84  
Db 285 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGATTCTGAGAGACAGGCTGGCGGCC 226  
Qy 85 TTAGATCTGTCAACAACGGTGTGATTACGGGTATTGCTAGCGCTTCCAAATGCTGGT 144  
Db 225 TTACGCCACCACACCCAGCTCGGCTGAGAGCTGAGCGAGCTGCTGCCAATGGTGGT 166  
Qy 145 GTGAATTTCTGAAGG 159  
Db 165 GGTCAATCAAGGAAGG 151

## RESULT 8

US-09-128-026-97/c  
; Sequence 97, Application US/09128026  
; Patent No. 6403335  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; MILBRANDT, JEFFREY D.  
; KOTZBAUER, PAUL T.

; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/128,026  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION/DOCKET NUMBER: 35,197  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-128-026-97

Query Match 18.2%; Score 32.6; DB 3; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.03;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 25 CACTTGATTCTTCGACAATACCTACCTGTGCGCCCTTTGGGAGATAATCCGACTGCA 84  
DB 285 CACTTGAGCTTTCAAAGGCTGCTTGTGATTCTGAGAGACAGGCTGGCGGCC 226  
QY 85 TTAAGATCTGTGAGAACACAGGTGTGGATTACGGGTATTGCTACGCCCTTCCAATGCTGGT 144  
DB 225 TTACGCCACCACAGCCACAAAGCTGCGGCTGAGAGCTGAGCAGCTGCTGCAATGGTGGT 166  
QY 145 GTGAATTTCTGAAGG 159  
DB 165 GGTCAATCAAGGAAGG 151

RESULT 9  
US-09-128-026-106/c  
; Sequence 106, Application US/09128026  
; Patent No. 6403335  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRANDT, JEFFREY D.  
; APPLICANT: KOTZBAUER, PAUL T.  
; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/128,026  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION/DOCKET NUMBER: 35,197  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 106:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-128-026-106

Query Match 18.2%; Score 32.6; DB 3; Length 336;  
Best Local Similarity 52.6%; Pred. No. 0.03;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 25 CACTTGATTCTTCGACAATACCTACCTGTGCGCCCTTTGGGAGATAATCCGACTGCA 84  
DB 285 CACTTGAGCTTTCAAAGGCTGCTTGTGATTCTGAGAGACAGGCTGGCGGCC 226  
QY 85 TTAAGATCTGTGAGAACACAGGTGTGGATTACGGGTATTGCTACGCCCTTCCAATGCTGGT 144  
DB 225 TTACGCCACCACAGCCACAAAGCTGCGGCTGAGAGCTGAGCAGCTGCTGCAATGGTGGT 166  
QY 145 GTGAATTTCTGAAGG 159  
DB 165 GGTCAATCAAGGAAGG 151

RESULT 10  
US-09-220-616-97/c  
; Sequence 97, Application US/09220616  
; Patent No. 6645937  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; APPLICANT: MILBRANDT, JEFFREY D.  
; APPLICANT: KOTZBAUER, PAUL T.  
; APPLICANT: LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: US  
; ZIP: 63105-1817  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/220,616  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981,739  
; FILING DATE: 31-Aug-1998  
; APPLICATION NUMBER: PCT/US97/03461  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.



```
;
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 976163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-220-616-97

Query Match      18.2%; Score 32.6; DB 4; Length 336;
Best Local Similarity 52.6%; Pred. No. 0.03;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 25 CACTTGATCTTCGACAAATACCTGTGCGGCCCTTTGGGAGATAATCCGAGCTGCA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGTGATTCTGAGAGACAGGCTGGCGCC 226

Qy 85 TTAAGATCTGTGCAAAACACGGTGTGATTACGGGTATTGTACGCCCTTCCAATGCTGGT 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 TTACGCCACACACCCACAGCTCGGCTGAGACTGAGGAGCTGTGCTGCAATGCTGGT 166

Qy 145 GTGAATTTCTGAAGG 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 GGTCAATCAAGGAAGG 151

RESULT 11
US-09-220-616-106/c
; Sequence 106, Application US/09220616
; Patent No. 6645937
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/220.616
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/981.739
; FILING DATE: 31-Aug-1998
; APPLICATION NUMBER: PCT/US97/03461
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 976163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-220-617-978-14

Query Match      18.2%; Score 32.6; DB 4; Length 336;
Best Local Similarity 52.6%; Pred. No. 0.03;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 25 CACTTGATCTTCGACAAATACCTGTGCGGCCCTTTGGGAGATAATCCGAGCTGCA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 CACCTGAGCCTTTCAAAGGCTGCTCTTGTGTGATTCTGAGAGACAGGCTGGCGCC 226

Qy 85 TTAAGATCTGTGCAAAACACGGTGTGATTACGGGTATTGTACGCCCTTCCAATGCTGGT 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 225 TTACGCCACACACCCACAGCTCGGCTGAGACTGAGGAGCTGTGCTGCAATGCTGGT 166

Qy 145 GTGAATTTCTGAAGG 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 GGTCAATCAAGGAAGG 151

RESULT 12
US-09-220-527-97/c
; Sequence 97, Application US/09220527
; Patent No. 6692943
; GENERAL INFORMATION:
; APPLICANT: JOHNSON JR., EUGENE M.
; APPLICANT: MILBRANDT, JEFFREY D.
; APPLICANT: KOTZBAUER, PAUL T.
; APPLICANT: LAMPE, PATRICIA A.
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTORS
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: US
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/220.527
; FILING DATE: 24-Dec-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/981.739
; FILING DATE: 31-Aug-1998
; APPLICATION NUMBER: PCT/US97/03461
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 976163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-220-527-97

Query Match      18.2%; Score 32.6; DB 4; Length 336;
```

	Best Local Similarity	52.6%;	Pred. No. 0.03;						
	Matches	71;	Conservative	0;	Mismatches	64;	Indels	0;	Gaps
Qy	25	CAC	TGATTCTTCG	CAGAA	TACCTG	TGCGCCCTTTCGG	GAGATAATCGG	ACTGCA	84
Db	285	CAC	TGAGCCTTTC	AAAGG	TGCCTT	TGCTTGTGATTCT	GAGAC	CAGGTCGCGCC	226
Qy	85	TTA	AGATCTGT	CAGAA	CACGGT	GTGAGTTACGGG	TATTGCTAC	GCTTCCAATGCTGGT	144
Db	225	TT	CAGCACCAC	AGCCACA	AGCTCGG	TGAGAGTGAGG	CAGCTGCTGCCAATGCTGGT	166	
Qy	145	GT	GAA	TTTCTG	AAGG				159
Db	165	GGT	CATCAAGG	AAGG					151

RESULT 13  
US-09-220-527-106/c  
; Sequence 106, Application US/09220527  
; Patent No. 6692943  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON JR., EUGENE M.  
; MILBRANDT, JEFFREY D.  
; KOTZBAUER, PAUL T.  
; LAMPE, PATRICIA A.  
; TITLE OF INVENTION: PERSPHIN AND RELATED GROWTH FACTORS

	Query Match	18.2%	Score 32.6;	DB 4;	Length 336;
	Best Local Similarity	52.6%;	Prod. No. 0.03;		
	Matches	71;	Conservative	0; Mismatches 64;	Indels 0; Gaps 0
Qy	25	CAC	TGTGATCTTCTTCGACAATACCTGTGCGGCCCTTTGGGAGATATATCGGACTGCA	84	
Db	285	CAC	TGAGCCCTTTCAGAGGCTGCTCTTGTCTGTGATTCTGAGAGACCAGGTGTGCCGCC	226	

Qy	85	TTAAGTCTGTGAGAACACGGTGTGGATTACGGGTATTGCTACGCTTCCAATGCTGGT	144
Db	225	TTGAGCACACAGCCACGAAGCTGCGGCTGAGAGCTGAGGACGCTGCTGCCAATGGTGGT	166
Qy	145	GTGAATTTCTGAAGG	159
Db	165	GGTCATCAAGGAAGG	151

RESULT 14  
US-09-220-407-97/c  
Sequence 97, Application US/09220407  
Patent No. 6716600  
GENERAL INFORMATION:  
APPLICANT: JOHNSON, EUGENE M  
APPLICANT: MILBRANDT, JEFFREY D  
APPLICANT: KOTZBAUER, PAUL T  
APPLICANT: LAMPE, PATRICIA A  
APPLICANT: KLEIN, ROBERT  
APPLICANT: DESAUVAGE, FRED  
TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
NUMBER OF SEQUENCES: 239  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORTSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MO  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/220,407  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/931,858  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 971486  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 97:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 336 base pairs  
Type: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-09-220-407-97

	Query Match	18.2%	Score 32.6;	DB 4;	Length 336;
	Best Local Similarity	52.6%;	Pred. No. 0.03;		
	Matches	71;	Conservative	0;	Mismatches 64; Indels 0; Gaps 0;
Qy	25	CACCTGATTCCTCCGACAATACCTACTCTGCGCCCTTTGGGAGATAATCCGGACCTGCA	84		
Db	285	CACCTGAGCCTTTCAAAGGCTGCCTCTTGCTTGATTTTGAGAGACCGCTGGCGGCC	226		
Qy	85	TTAAGATCTGTCAAGAACACGGTGTGGATTACGGGTATTGCTACGGCCTTCCAAATGCTGGT	144		
Db	225	TTCAGCCACACAGCCACAGCTGCGGCTGAGAGCTGAGGCAGCTGCTGCCAATGCTGGT	166		
Qy	145	GTGAATTTCTGAAGG	159		
Db	165	GGTCATCAAGCAAGG	151		

## RESULT 15

US-08-931-858E-107/c  
; Sequence 107, Application US/08931858E  
; Patent No. 622022  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, EUGENE M  
; APPLICANT: MILBRANDT, JEFFREY D  
; APPLICANT: KOTZBAUER, PAUL T  
; APPLICANT: LAMPE, PATRICIA A  
; APPLICANT: KLEIN, ROBERT  
; APPLICANT: DESAUVAGE, FRED  
; TITLE OF INVENTION: PERSEPHIN AND RELATED GROWTH FACTOR  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/931.858E  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971486  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 391 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; US-08-931-858E-107

Query Match 18.2%; Score 32.6; DB 3; Length 391;  
Best Local Similarity 52.6%; Pred. No. 0.032; 64; Indels 0; Gaps 0;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
  
Qy 25 CACTTGATTCTTCGACAATACCTGTGCGCCCTTTGGGAGATAATCCGGACTGCA 84  
Db 337 CACCTGAGCCTTTCAGAGGCTGCTCTTGTGATTCTGAGAGACGAGCTGCGCC 278  
  
Qy 85 TTAAGATCTGTGAGAAACACGGTGTGATTACGGGTATTGTACGCCCTTCAATGCTGGT 144  
Db 277 TTCAGCCACCACACGACAGCTCGGCTGAGAGCTGAGGAGCTGCTGCCAATGCTGGT 218  
  
Qy 145 GTGAATTCTGAAGS 159  
Db 217 GGTATCAAGGAGG 203

Search completed: July 30, 2005, 20:19:10  
Job time : 87.4775 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:50:29 ; Search time 1960.96 Seconds  
(without alignments)  
3474.585 Million cell updates/sec

Title: US-10-617-978-14\_COPY\_62\_240  
Perfect score: 179  
Sequence: 1 cgctgacgtcccggaac.....atgagaacgtgaaggtctga 179

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	35	19.6	414	2	AW343516	f145a12.x
C 2	34.6	19.3	728	8	B2047564	B2047564 lkh84c10.
C 3	33.8	18.9	411	6	CB803038	AMGNNUC:M
C 4	33.8	18.9	595	8	B2506790	BONHL22TR
C 5	33.8	18.9	857	8	BH535763	BOGQH46TF
C 6	33.6	18.8	1033	9	CNS013FB	Drosophil
C 7	33.2	18.5	1092	7	CN063005	Ag2 p39 O
C 8	33	18.4	580	8	BH741832	gt38b03.g
C 9	33	18.4	644	8	BH675875	BOMDB38TR
C 10	33	18.4	788	8	BH439076	BOGJM34TR
C 11	33	18.4	835	8	BZ501283	BONLH48TR
C 12	32.8	18.3	637	7	C0686621	DG11-2199
C 13	32.6	18.2	426	7	C0141963	EST838634
C 14	32.6	18.2	454	7	C0143971	EST838642
C 15	32.6	18.2	455	7	C0134628	EST829299
C 16	32.6	18.2	744	9	CL172321	104_374_1
C 17	32.2	18.0	609	7	C0584797	DG2-1081
C 18	32.2	18.0	784	8	BZ444956	BONJU06TF
C 19	32	17.9	550	2	BE767482	RCO-NT012
C 20	32	17.9	1101	9	CNS014SJ	Drosophil
C 21	31.8	17.8	484	5	B0300487	PMO-KT004
C 22	31.8	17.8	543	2	BF770477	RC2-IT004
C 23	31.8	17.8	543	2	BF771057	RC2-IT004
C 24	31.8	17.8	546	2	BF770481	RC2-IT004

25	31.8	17.8	546	2	BF771074	BF771074 RC2-IT004
26	31.8	17.8	656	7	CV092608	PMU USDA
27	31.6	17.7	282	1	AV341134	AV341134
28	31.6	17.7	417	4	B1880995	f83b12.Y
C 29	31.6	17.7	713	8	BZ456120	BONJES7TR
C 30	31.6	17.7	758	7	CK027614	AGENCOURT
C 31	31.6	17.7	794	8	BH486210	BOGUO61TR
C 32	31.6	17.7	807	8	BZ447705	BZ447705 BONRY58TF
C 33	31.6	17.7	875	5	BX338057	BX338057
C 34	31.6	17.7	886	7	CNS05232	AGENCOURT
C 35	31.6	17.7	896	6	CD253204	AGENCOURT
C 36	31.6	17.7	960	1	AL517987	AL517987
C 37	31.4	17.5	466	4	BI673093	ft34b05.Y
C 38	31.4	17.5	476	8	BZ175692	CH230-397
C 39	31.4	17.5	573	8	BZ302592	Forward s
C 40	31.4	17.5	577	9	CR249625	Forward s
C 41	31.4	17.5	637	8	BH578500	BOHC174TF
C 42	31.4	17.5	669	8	AZ574720	331PvH05
C 43	31.4	17.5	678	8	BH512431	BOGQY14TF
C 44	31.4	17.5	689	8	BH950118	od195R04.
C 45	31.4	17.5	717	8	BH548329	BOHLM46TR

## ALIGNMENTS

RESULT 1  
AW343516/c  
LOCUS  
DEFINITION  
f145a12.x1 Sugano Kawakami zebrafish DNA Danio rerio cDNA clone  
INAGS:2640574.3, similar to SW:MCM2\_XENLA P55961 DNA REPLICATION  
LICENSING FACTOR MCM2 ;, mRNA sequence.  
ACCESSION  
AW343516  
VERSION  
AW343516.1 GI:68399882  
KEYWORDS  
EST.  
SOURCE  
Danio rerio (zebrafish)  
ORGANISM  
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE  
1 (bases 1 to 414)  
Sugano S., Kawakami K., Johnson S., Li F., Marra M., Eddy S.,  
Hillier L., Clifton S., Allen M., Gibbons M., Jost S., Kucaba T.,  
Martin J., Pape D., Steptoe M., Underwood K., Theising B.,  
Ritter E., Bowers Y., Wyllie T., Waterston R. and Wilson R.  
WashU Zebrafish EST Project 1999  
Unpublished (1999)  
CONTACT: S.L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA  
Sequencing by: Washington University Genome Sequencing Center  
Seq primer: T7 ET from Amersham  
High quality sequence stop: 163.  
FEATURES  
Location/Qualifiers  
1..414  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/cldones="IMAGE:2640574"  
/sex="mixed (one male and one female, including  
unfertilized eggs)"  
/dev\_stage="adult"  
/clone\_host="Dn10B (phage resistant)"  
/clone\_lib="Sugano Kawakami zebrafish DRH"  
/note="Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG);  
Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed  
with an oligo(dT) primer (ATCGCCCTTTTCTTTTCTTT);  
double-stranded cDNA was ligated to a DraIII adaptor



```

Seq primer: TF
Class: sheared ends.
FEATURES             source
    Location/Qualifiers
        1..857
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="TOL000DH3"
            /db_xref="taxon:3712"
            /clone_lib="BOGQH46"
            /clone_lib="BOGQ"
            /note="Vector: pHOS1; Site 1: BatXI; 2-3 kb sheared
            genomic DNA inserted into pHOS1 using BatXI linkers"

ORIGIN
Query Match          18.9%; Score 33.8; DB 8; Length 857;
Best Local Similarity 53.4%; Pred. No. 3.9;
Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      28  TTGATCTTCGACACATACCTACCTGTGCGCCCTTGGAGATATCCGACTGCATTA 87
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Ddb     50  TTGCTTATATAGAAAATAGTTCTTGTTGCGCGGTTCTTCCTGTATCAGGATGCCAC 109
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      88  AGATCTGTCAGAAACACGGTGGGATTACGGGTATTGTACGCCCTTCCCAATGCTGGTGTG 147
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Ddb     110 ACAGCTGTGTCGACACGGCGATGTACAGATAAAGTTTCCCTCCACGGGTTTACG 169
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      148 AATTTCTGAAGCA 160
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Ddb     170 AGGACTGGAGGGA 182
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
CNS013FB
LOCUS
DEFINITION
  CNS013FB 1033 bp DNA linear GSS 26-JUL-1999
  Drosophila melanogaster genome survey sequence T7 end of BAC
  BACN09J03 of DrosBAC library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
  AL102785
  AL102785.1 GI:5614396
  GSS.
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 1033)
  Genoscope.
  Direct Submission
  Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (EDGP) -
  http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
  library (Dros BAC) was made by Alain Billaud at CEPH (Centre
  d'Etude du Polymorphisme Humain) with funding provided by a MRC
  project grant. The DNA was prepared from embryos by Alain Bucheton
  and Genevieve Payan. It has been constructed in the vector
  pBelOBAC11.

FEATURES             source
    Location/Qualifiers
        1..1033
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN09J03"
            /clone_lib="DrosBAC"
            /plasmid="pBelOBAC11"
            /note="end : T7"

ORIGIN
Query Match          18.8%; Score 33.6; DB 9; Length 1033;
Best Local Similarity 23.9%; Pred. No. 4.8;
Matches 34; Conservative 43; Mismatches 65; Indels 0; Gaps 0;

```

```
QY 17 AAACACTCCCACTTGATCTTCCGACAATACTACCTGTGCGCCCTTTGGAGATATCC 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 892 MAMNNMNNMNNMNTTNTTNTTNNMNNMNNMNNMNNMNNMNTTVGGGRTTMMN 951
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 GGACTGCATTAAGATCTGTGAGAAACACGGTGTGGATTACGGGTATTGCTACGCCCTTCCA 136
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 952 GMRGTGTTTMMMGMGTCGCMGNTVTMMGGGTRTCTKMTTGTTRTKTDKTCGCMGNTGM 1011
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 ATGCTGGTGTGAATTTCTCTGAAG 158
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1012 KMGMGTGTRKTGTGKMGMGMK 1033
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
LOCUS CN063005 1092 bp mRNA linear EST 30-MAR-2004
DEFINITION Ag2_p39_024_M13R AG Ambystoma tigrinum tigrinum cDNA, mRNA
ACCESSION CN063005
VERSION CN063005.1 GI:45839068
KEYWORDS EST.
SOURCE Ambystoma tigrinum tigrinum (Eastern tiger salamander)
ORGANISM Ambystoma tigrinum tigrinum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
Ambystoma.
REFERENCE 1 (bases 1 to 1092)
AUTHORS Putta,S., Smith,J.J., Walker,J.A., Rondet,M., Weisrock,D.,
Monaghan,J., Samuels,A.K., Kump,K., King,D.C., Maness,N.J.,
Habermann,B., Tanaka,E., Bryant,S.V., Gardiner,D.M., Parichy,D.M.
and Voss,S.R.
TITLE From biomedicine to natural history research: EST resources for
ambystomatid salamanders
JOURNAL BMC Genomics 5 (1), 54 (2004)
COMMENT Contact: SR Voss
Department of Biology
University of Kentucky
TH Morgan Building, Lexington, KY 40506, USA
Tel: 859 257 9888
Fax: 859 257 1717
Email: srvooss@uky.edu
The EST is quality trimmed at the ends with a 20 base window and
quality threshold of 15 (phred quality score). Please visit
http://salamander.uky.edu For any information(trace,quality files
etc) regarding this EST.
FEATURES
source
location/Qualifiers
1..1092
/organism="Ambystoma tigrinum tigrinum"
/mol_type="mRNA"
/sub_species="tigrinum"
/db_xref="taxon:43116"
/tissue_type="Liver, Lung, Kidney, Heart, gonad, brain and
gill tissues collected from metamorphosing larvae"
/clone_lib="AG"

ORIGIN
Query Match 18.5%; Score 33.2; DB 7; Length 1092;
Best Local Similarity 51.3%; Pred. No. 6.7;
Matches 77; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 29 TGATCTTCCGACAATACTACCTGTGCGCCCTTTGGGAGATAATCCGGACTGCATTAA 88
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 99 TGAGTTAGATGACGATAGAGAAGAAATGCATCGTTGTGACTATTGTGCGTGGGAACCA 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89 GATCTGTGAGAAACACCGGTGTGGATTACGGGTATTGTACGCCCTTCCAAATGCTGGTGA 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 TGGTTGTGACGATGAGTGTGTCAACACCGCAAGCTTACGCTCTGCAAAATGCGGCCAAGG 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 ATTCTTGAGGATGAGAACGCTGAGGTCG 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 ATTCTTCTCAATCCAGACCTGAAGACCTG 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 8
LOCUS BH741832/c 580 bp DNA linear GSS 25-FEB-2002
DEFINITION gt38b03.g1 BoBuds01 Brassica oleracea genomic clone gt38b03 5',
genomic survey sequence.
ACCESSION BH741832
VERSION BH741832.1 GI:18876445
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 580)
AUTHORS Katari,M., O'Shaughnessy,A., Palmer,L., Bahret,A., Baker,J.,
Bailja,V., Cummins,D.M., Katzenberger,F., King,L., Kirchoff,K.,
Kuit,K., Miller,B., Muller,S., Nascimento,L., Preston,R.,
Santos,L., Shah,R., Zutavern,T., Dedhia,N., Rabinowicz,P.D. and
McCombie,W.R.
TITLE Whole Genome Shotgun Reads from Brassica oleracea (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: gt38 row: b column: 03
Seq primer: -21UnivRev
Class: shotgun
High quality sequence stop: 580.
FEATURES
source
location/Qualifiers
1..580
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="gt38b03"
/clone_lib="BoBuds01"
/note="Vector: M13 for .x reads, pBluescript for .b and .g
reads; Site 1: EcoRV; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear prep
using Brassica oleracea T01000DH3 buds provided by Thomas
Osborn at the University of Wisconsin. Genomic DNA
provided by Pablo Rabinowicz (CSHL) and shotgun library
prepared in McCombie Lab."

ORIGIN
Query Match 18.4%; Score 33; DB 8; Length 580;
Best Local Similarity 54.5%; Pred. No. 6.6;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 28 TTGATCTTCCGACAATACTACCTGTGCGCCCTTTGGGAGATAATCCGGACTGCATTA 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 TTGCTTATATAGAAAATAGATTCTTTCGCCGCGTCTTCTCCCTGATCAGGACGCGCTC 440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 88 AGATCTGTGAGAAACACGGTGTGGATTACGGGTATTGCTACGCCCTTCCAATGCTGGTGC 147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 ACAGCTGTGTTGCCAACACACGCGATGTACAAGAAACAAGGTTCCTCCCTCCACGGGTTTTCG 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 148 A 148
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 A 379

RESULT 9
LOCUS BH675875 644 bp DNA linear GSS 19-FEB-2002
DEFINITION BOMDB38TR BO_2_3_KB Brassica oleracea genomic clone BOMDB38,
genomic survey sequence.
ACCESSION BH675875
VERSION BH675875.1 GI:18746318
```





```

Qy 28 TTGATTCTTCGCAAAATACCTACTGTGCGCCCTTTGGGAGATAATCCGAGCTGCATTA 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 TTGCTTATATAGAAAATAGGTTTCTGTTCGCGGTTCTGTCCTCTGATCAGGACGCCCTC 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 88 AGATCTCTCGAAACACGGTGTGATACGGGTATTGCTAGCCCTTCCATGCTGGTGTG 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 ACAGCTGTTCGCGACACGGGACGTACAGAAAGGTTTCCCTCCACAGGTTTTCG 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 148 A 148
Db 283 A 283

RESULT 12
CO686621/c
LOCUS
DEFINITION DG11-219g11 DG11-kidney Canis familiaris cDNA 3', mRNA sequence.
ACCESSION CO686621
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 637)
AUTHORS Schluter,T., Hermanns,J., Weindel,M., Schuette,D., Kranz,H.,
Henrich,J. and Loebbert,R.
TITLE Dog arrayTAG cDNA clone collection
JOURNAL Unpublished (2004)
COMMENT Contact: Thomas Schluter
LION bioscience AG
Waldhoferstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schluter@lionbioscience.com.

FEATURES
    source
        1..637
            /organism="Canis familiaris"
            /mol_type="mRNA"
            /strain="Beagle"
            /db_xref="taxon:9615"
            /tissue_type="kidney"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="DG11-kidney"
            /note="Organ: kidney; Vector: Dog pBluescript LION"

ORIGIN
Query Match 18.3%; Score 32.8; DB 7; Length 637;
Best Local Similarity 53.0%; Pred. No. 7.9;
Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 31 ATTCTTCGGAATACCTACTGTGCGCCCTTTGGGAGATAATCCGAGCTGCATTAAGA 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 ATTCACAAAACAGACAGCTCCTGCTGTACTGCTGGATTTTCAGGCACTGCATTATA 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 91 TCTGTCAAGAACACGGGTGTGATACGGGTATTCTACGCTTCCATGCTGCTGAAT 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 ACCTTCTGAACAAATGTAGTTTCCCTATTGGGACAAATCTATTCTTGTGACGT 25
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 151 TTCTGAAGGATG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24 TATTAAATCTTG 13

RESULT 13
CO141963
LOCUS
DEFINITION EST836634 Aspergillus flavus Normalized cDNA Expression Library
ACCESSION CO141963
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Aspergillus flavus
Aspergillus; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 426)
AUTHORS Yu,J., Whitelaw,C.A., Nierman,W.C., Bhatnagar,D. and Cleveland,T.E.
TITLE Aspergillus flavus expressed sequence tags for identification of
genes with putative roles in aflatoxin contamination of crops
JOURNAL PEMS Microbiol. Lett. (2004) In press
COMMENT Contact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jiu@ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRRC (jiu@ars.usda.gov) for clone
information
PCR Primers
FORWARD: M13F
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES
    source
        1..426
            /organism="Aspergillus flavus"
            /mol_type="mRNA"
            /strain="NRRL 3357"
            /db_xref="taxon:5059"
            /clone="NAFEJ04"
            /sex="asexual mycelia"
            /cell_type="mycelia"
            /dev_stage="developmental stages from 18 to 96 hours"
            /lab_host="E. coli DH10B T1 resistant cells"
            /clone_lib="Aspergillus flavus Normalized cDNA Expression
            Library"
            /note="Vector: pBlueScript (SK+) (Stratagene), antibiotic
            selection marker: Carbenicillin; Site 1: NotI, at the 5
            prime end; Site 2: EcoRI, at the 3 prime end; This
            normalized cDNA expression library was constructed using a
            mixture of mycelial cells grown under eight different
            medium conditions and harvested at 5 time points (18, 24,
            48, 72, 96 hours). The poly-A sequence was trimmed off
            before ligating to vector."

ORIGIN
Query Match 18.2%; Score 32.6; DB 7; Length 426;
Best Local Similarity 52.6%; Pred. No. 8.3;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 37 CCGACAAATACCTACTGTGCGCCCTTTGGGAGATAATCCGAGCTGCATTAGATCTGTC 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 CCGTCACAAACGCCCGTGACGCATTTATCCGTGAGCAATGGGTCCGACGATGATCTTC 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCCCTTCCAATGCTGGTGAATTTCTGA 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 GTCTCTCGTGTGAGGAAATTTGGGGAAGTGTATGCGCGAGGCGGTGAATCTTGGAGA 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 157 AGGATGAGAACGTGA 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 AGTGTGGGCTTTGA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
CO143971
LOCUS
DEFINITION EST838642 Aspergillus flavus Normalized cDNA Expression Library
ACCESSION CO143971
VERSION CO143971.1 GI:48899792
KEYWORDS EST.
SOURCE Aspergillus flavus

```

ORGANISM Aspergillus flavus  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE 1 (bases 1 to 454)  
AUTHORS Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.  
TITLE Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops  
JOURNAL FEMS Microbiol. Lett. (2004) In press  
COMMENT Contact: Yu J  
Food and Feed Safety Research Unit  
USDA/ARS, Southern Regional Research Center  
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
Tel: 504 286 4405  
Fax: 504 286 4419  
Email: jiyu@rrc.ars.usda.gov  
Contact Dr. Yu at USDA/ARS SRRRC (jiyu@rrc.ars.usda.gov) for clone information  
PCR Primers  
FORWARD: M13F  
BACKWARD: M13R  
Seq primer: M13 Forward  
POLYA=No.

## FEATURES

source

Location/Qualifiers

1..454

/organism="Aspergillus flavus"  
/mol\_type="mRNA"  
/strain="NRRL 3357"  
/db\_xref="taxon:5059"  
/clone="NAFFH21"  
/sex="asexual mycelia"  
/cell\_type="mycelia"  
/dev\_stage="developmental stages from 18 to 96 hours"  
/lab\_host="E. coli DH10B T1 resistant cells"  
/clone\_lib="Aspergillus flavus Normalized cDNA Expression Library"

/note=vector: pBlueScript (SK+) (Stratagene), antibiotic selection marker; Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

## ORIGIN

Query Match 18.2%; Score 32.6; DB 7; Length 454;  
Best Local Similarity 52.6%; Pred. No. 8.4;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 37 CCGACAATACCTACTGTGCGCCCTTTGGGAGATAATCCGACTGCATTGAATCTGTC 96  
Db 25 CCGTCCACACGCCCGTGACGCCATTATCCGTGAGCAATGGGTCGCGACGATGATGTC 84

Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGAATTTCTGA 156  
Db 85 GTCTCGTGGTGAGGAATTTGGGGAAGTGCTATGCGCGAGGCGGTGAATCATTTGGAGA 144

Qy 157 AGGATGAGAACGTGA 171  
Db 145 AGTGTGGGGCTTTGA 159

## RESULT 15

LOCUS COL134628 455 bp mRNA linear EST 17-JUN-2004  
DEFINITION EST82299 Aspergillus flavus Normalized cDNA Expression Library  
Aspergillus flavus cDNA clone NAFU67 5' end, mRNA sequence.  
ACCESSION COL134628  
VERSION COL134628.1 GI:48883606  
KEYWORDS EST.  
SOURCE Aspergillus flavus  
ORGANISM Aspergillus flavus  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

1 (bases 1 to 455)  
Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.  
Aspergillus flavus expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops  
FEMS Microbiol. Lett. (2004) In press  
Contact: Yu J  
Food and Feed Safety Research Unit  
USDA/ARS, Southern Regional Research Center  
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
Tel: 504 286 4405  
Fax: 504 286 4419  
Email: jiyu@rrc.ars.usda.gov  
Contact Dr. Yu at USDA/ARS SRRRC (jiyu@rrc.ars.usda.gov) for clone information  
PCR Primers  
FORWARD: M13F  
BACKWARD: M13R  
Seq primer: M13 Forward  
POLYA=No.

## FEATURES

source

Location/Qualifiers

1..455

/organism="Aspergillus flavus"  
/mol\_type="mRNA"  
/strain="NRRL 3357"  
/db\_xref="taxon:5059"  
/clone="NAFAU67"  
/sex="asexual mycelia"  
/cell\_type="mycelia"  
/dev\_stage="developmental stages from 18 to 96 hours"  
/lab\_host="E. coli DH10B T1 resistant cells"  
/clone\_lib="Aspergillus flavus Normalized cDNA Expression Library"

/note=vector: pBlueScript (SK+) (Stratagene), antibiotic selection marker; Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

## ORIGIN

Query Match 18.2%; Score 32.6; DB 7; Length 455;  
Best Local Similarity 52.6%; Pred. No. 8.4;  
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 37 CCGACAATACCTACTGTGCGCCCTTTGGGAGATAATCCGACTGCATTGAATCTGTC 96  
Db 41 CCGTCCACACGCCCGTGACGCCATTATCCGTGAGCAATGGGTCGCGACGATGATGTC 100

Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGAATTTCTGA 156  
Db 101 GTCTCGTGGTGAGGAATTTGGGGAAGTGCTATGCGCGAGGCGGTGAATCATTTGGAGA 160

Qy 157 AGGATGAGAACGTGA 171  
Db 161 AGTGTGGGGCTTTGA 175

Search completed: July 30, 2005, 20:16:19  
Job time : 1968.96 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 18:20:14 ; Search time 385.323 Seconds  
(without alignments)  
2973.108 Million cell updates/sec

Title: US-10-617-978-17\_COPY\_73\_249

Perfect score: 177

Sequence: 1 gccagctgccagggaacta.....acgagaacgtcaaggtgtga 177

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_PUBCOMB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.8	24.7	2748	15	US-10-156-761-5636 Sequence 5636, Ap
2	43.8	24.7	9025608	15	US-10-156-761-1 Sequence 1, Appl
3	41.6	23.5	192	21	US-10-721-793-75 Sequence 75, Appl
4	41.6	23.5	316	21	US-10-721-793-73 Sequence 73, Appl
5	41.6	23.5	675	13	US-10-006-922-38 Sequence 38, Appl
6	41.6	23.5	675	13	US-10-081-864-13 Sequence 13, Appl
7	41.6	23.5	678	14	US-10-006-922-36 Sequence 36, Appl

8	41.6	23.5	678	14	US-10-081-864-7 Sequence 7, Appl
9	41.6	23.5	678	14	US-10-081-864-14 Sequence 14, Appl
10	41.6	23.5	678	14	US-10-121-258-5 Sequence 5, Appl
11	41.6	23.5	678	16	US-10-315-920-1 Sequence 1, Appl
12	41.6	23.5	678	16	US-10-315-920-3 Sequence 3, Appl
13	41.6	23.5	678	16	US-10-315-920-5 Sequence 5, Appl
14	41.6	23.5	681	13	US-10-006-922-35 Sequence 35, Appl
15	41.6	23.5	681	13	US-10-006-922-37 Sequence 37, Appl
16	41.6	23.5	681	14	US-10-121-258-3 Sequence 3, Appl
17	41.6	23.5	681	14	US-10-121-258-23 Sequence 23, Appl
18	41.6	23.5	704	22	US-10-844-064A-3 Sequence 3, Appl
19	41.6	23.5	723	14	US-10-152-296-1 Sequence 1, Appl
20	41.6	23.5	723	19	US-10-739-656-1 Sequence 1, Appl
21	41.6	23.5	747	20	US-10-785-862-10 Sequence 7, Appl
22	41.6	23.5	1050	13	US-10-060-857-7 Sequence 51, Appl
23	41.6	23.5	1638	15	US-10-214-932-51 Sequence 75, Appl
24	41.6	23.5	1647	15	US-10-214-932-75 Sequence 9, Appl
25	41.6	23.5	4200	22	US-10-894-949-9 Sequence 8, Appl
26	41.6	23.5	4300	22	US-10-894-949-8 Sequence 29, Appl
27	41.6	23.5	4692	15	US-10-161-403-29 Sequence 16, Appl
28	41.6	23.5	4692	19	US-10-433-640-16 Sequence 46, Appl
29	41.6	23.5	5436	21	US-10-169-050-46 Sequence 45, Appl
30	41.6	23.5	6984	13	US-10-001-189-45 Sequence 5, Appl
31	41.6	23.5	7495	19	US-10-742-828-5 Sequence 4, Appl
32	41.6	23.5	7508	19	US-10-742-828-4 Sequence 20, Appl
33	41.6	23.5	7910	21	US-10-169-050-20 Sequence 20, Appl
34	41.6	23.5	9320	19	US-10-471-065-20 Sequence 4, Appl
35	41.6	23.5	9658	19	US-10-609-019-4 Sequence 3, Appl
36	41.6	23.5	9678	19	US-10-609-019-3 Sequence 5, Appl
37	41.6	23.5	9731	22	US-10-510-363-5 Sequence 6, Appl
38	41.6	23.5	9782	22	US-10-510-363-6 Sequence 2, Appl
39	41.6	23.5	10263	19	US-10-609-019-2 Sequence 3, Appl
40	41	23.2	726	16	US-10-314-827-3 Sequence 1272, Ap
41	40.4	22.8	263	9	US-09-923-876-1272 Sequence 1272, Ap
42	40.4	22.8	263	10	US-09-923-876-1272 Sequence 5616, Ap
43	40.4	22.8	714	17	US-10-260-238-5616 Sequence 13478, A
44	40.4	22.8	1042	18	US-10-425-114-13478 Sequence 20725, A
45	40.4	22.8	1131	18	US-10-425-114-20725

#### ALIGNMENTS

RESULT 1  
US-10-156-761-5636  
; Sequence 5636, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10156.761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 5636  
; LENGTH: 2748  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2748)  
US-10-156-761-5636

```
Query Match      24.7%; Score 43.8; DB 15; Length 2748;
Best Local Similarity 60.5%; Pred. No. 0.0007;
Matches 72; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 48 CTTGTGGCACCCTGTGGCGACAAACCCGGAGTGCATCAAGATCTGCCAGAGCAGCGCGT 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 CGTGTGGCGGAGCAGCGCGCGACCCGGAGTGCCTTCTTCCACGAGGTTCGGCCT 2657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 108 CGACTACGGCTACTGCTACCGGTTCCAGTGTGGTGGAGTTCCTGAAGGACGAGAACG 166
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2658 CGACTACGTGCTCTGCTCGCGGTTCCGGATTCCGGTGGCCCGCTGGAGGCGGGCGCG 2716
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match      24.7%; Score 43.8; DB 15; Length 9025608;
Best Local Similarity 60.5%; Pred. No. 0.0018;
Matches 72; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 48 CTTGTGGCACCCTGTGGCGACAAACCCGGAGTGCATCAAGATCTGCCAGAGCAGCGCGT 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6840426 CGTGTGGCGGAGCAGCGCGCGACCCGGAGTCCGTCCTTCTTCCACGAGGTTCGGCCT 6840485
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 108 CGACTACGGCTACTGCTACCGGTTCCAGTGTGGTGGAGTTCCTGAAGGACGAGAACG 166
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6840486 CGACTACGTGCTCTGCTCGCGGTTCCGGATTCCGGTGGCCCGCTGGAGGCGGGCGCG 6840544
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-10-721-793-75
; Sequence 75, Application US/10721793
; Publication No. US2005006531A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
```

```
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Centruroides limpidus limpidus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(192)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
US-10-721-793-75

Query Match      23.5%; Score 41.6; DB 21; Length 192;
Best Local Similarity 60.7%; Pred. No. 0.0026;
Matches 68; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 35 CGGACAAACACCTACCTGTGGCACCCTGTGGCGACAAACCCGGAGTGCATCAAGATCTGCC 94
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 CGGGCTGCAAAATACGGGTGCTATGAATTGGGTGACAAACGGTTACTCGGATAGGAATGCA 88
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 95 AGAAGCACGGCGTCGACTAGCGCTACTGCTACCGGTTCCAGTTCAGTTCGTCGCA 146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 AAGCGGAGAGCGGTAACCTACGGCTATTGCTATATACTGTTGGGTGCTGTCGTCGCA 140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-10-721-793-73
; Sequence 73, Application US/10721793
; Publication No. US2005006531A1
; GENERAL INFORMATION:
; APPLICANT: Corona Villegas, Miguel
; APPLICANT: Garcia Rodriguez, Ma Consuelo
; APPLICANT: Valdez Cruz, Norma Adriana
; APPLICANT: Gurrola Briones, Georgina
; APPLICANT: Becerril Lujan, Baltazar
; APPLICANT: Possani Postay, Lourival Domingos
; TITLE OF INVENTION: Recombinant Immunogens for the Generation of Antivenoms to the
; FILE REFERENCE: 2099.0070001
; CURRENT APPLICATION NUMBER: US/10/721,793
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: US 60/430,067
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Centruroides limpidus limpidus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(259)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin precursor
; OTHER INFORMATION: In the mature peptide, the last Ser is amidated, and the last Gly
; OTHER INFORMATION: and the last basic aminoacid are cut
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (263)..(316)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(4)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (62)..(1)
; OTHER INFORMATION: Product= Sodium-channel modifier toxin
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (5)..(61)
; OTHER INFORMATION:
```



## US-10-006-922-36

Query Match 23.5%; Score 41.6; DB 13; Length 678;  
Best Local Similarity 53.0%; Pred. No. 0.003;  
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 2 CCGACGTGCCAGGGAACACTACCCGCTGGACAGCTCGGACAAACCTACCTGTGGCGACCCC 61  
DB 206 CCAAGGTGTACGTGAAGCACCCTCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 265  
QY 62 TGGCGCACACCCGGACTGCATCAAGATCTGCCAGAGCAGCGGCTGCAGTACGGCTACT 121  
DB 266 AGGCTTCAAGTGGAGCGCGTGTGAATCTCGAGGACGCGCGTGGCGACCGTGACCC 325  
QY 122 GCTACCGTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169  
DB 326 AGGACTCTCTCCCTGCAGGACGGCTGCTTCACTACAAAGTGAAGTTCA 373

## RESULT 8

US-10-081-864-7  
; Sequence 7, Application US/10081864  
; Publication No. US20030022287A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Savitsky, Alexandr  
; APPLICANT: Fradkov, Arcady  
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
; FILE OF INVENTION: Methods for Using the Same  
; CURRENT APPLICATION NUMBER: US/10/081,864  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/270,983  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp  
US-10-081-864-7

Query Match 23.5%; Score 41.6; DB 14; Length 678;  
Best Local Similarity 53.0%; Pred. No. 0.003;  
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 2 CCGACGTGCCAGGGAACACTACCCGCTGGACAGCTCGGACAAACCTACCTGTGGCGACCCC 61  
DB 206 CCAAGGTGTACGTGAAGCACCCTCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 265  
QY 62 TGGCGCACACCCGGACTGCATCAAGATCTGCCAGAGCAGCGGCTGCAGTACGGCTACT 121  
DB 266 AGGCTTCAAGTGGAGCGCGTGTGAATCTCGAGGACGCGCGTGGTGCACCGTGACCC 325  
QY 122 GCTACCGTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169  
DB 326 AGGACTCTCTCCCTGCAGGACGGCTGCTTCACTACAAAGTGAAGTTCA 373

## RESULT 9

US-10-081-864-14  
; Sequence 14, Application US/10081864  
; Publication No. US20030022287A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Savitsky, Alexandr  
; APPLICANT: Fradkov, Arcady  
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
; FILE OF INVENTION: Methods for Using the Same  
; CURRENT APPLICATION NUMBER: US/10/081,864  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/270,983  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
US-10-081-864-14

Query Match 23.5%; Score 41.6; DB 14; Length 678;  
Best Local Similarity 53.0%; Pred. No. 0.003;  
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 2 CCGACGTGCCAGGGAACACTACCCGCTGGACAGCTCGGACAAACCTACCTGTGGCGACCCC 61  
DB 206 CCAAGGTGTACGTGAAGCACCCTCCGCGACATCCCGGACTACAAGAGCTGTCTTCCCG 265  
QY 62 TGGCGCACACCCGGACTGCATCAAGATCTGCCAGAGCAGCGGCTGCAGTACGGCTACT 121  
DB 266 AGGCTTCAAGTGGAGCGCGTGTGAATCTCGAGGACGCGCGTGGCGACCGTGACCC 325  
QY 122 GCTACCGTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169  
DB 326 AGGACTCTCTCCCTGCAGGACGGCTGCTTCACTACAAAGTGAAGTTCA 373

## RESULT 10

US-10-121-258-5  
; Sequence 5, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
US-10-121-258-5

## RESULT 10

US-10-121-258-5  
; Sequence 5, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
US-10-121-258-5

## RESULT 10

US-10-121-258-5  
; Sequence 5, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
US-10-121-258-5



```
Db 266 AGGCTTCAAGTGGAGCGGTGATGAACCTTCGAGGACGGCGGTGGTGAACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTCTTCACTCAAGGTGAAGTTCA 373

RESULT 11
US-10-315-920-1
; Sequence 1, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-1
Query Match 23.5%; Score 41.6; DB 16; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACCTACTGTCGCGACCCC 61
Db 206 CCAAGGTGTACGTGAAGCACCACCCCGCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 265
Qy 62 TGGCGGACAAACCCGGAGCTGCATCAAGATCTCCAGAGACGCGGTGCACTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGCGTGTATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTTGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTGCTTCATCTACAGGTGAAGTTCA 373

RESULT 12
US-10-315-920-3
; Sequence 3, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-3
Query Match 23.5%; Score 41.6; DB 16; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACCTACTGTCGCGACCCC 61
Db 206 CCAAGGTGTACGTGAAGCACCACCCCGCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 265
Qy 62 TGGCGGACAAACCCGGAGCTGCATCAAGATCTCCAGAGACGCGGTGCACTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGCGTGTATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTTGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTGCTTCATCTACAGGTGAAGTTCA 373

RESULT 13
US-10-315-920-5
; Sequence 5, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-5
Query Match 23.5%; Score 41.6; DB 16; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACCTACTGTCGCGACCCC 61
Db 206 CCAAGGTGTACGTGAAGCACCACCCCGCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 265
Qy 62 TGGCGGACAAACCCGGAGCTGCATCAAGATCTCCAGAGACGCGGTGCACTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGCGTGTATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTTGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTGCTTCATCTACAGGTGAAGTTCA 373

RESULT 14
US-10-006-922-35
; Sequence 35, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
```

```
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-3
Query Match 23.5%; Score 41.6; DB 16; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACCTACTGTCGCGACCCC 61
Db 206 CCAAGGTGTACGTGAAGCACCACCCCGCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 265
Qy 62 TGGCGGACAAACCCGGAGCTGCATCAAGATCTCCAGAGACGCGGTGCACTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGCGTGTATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTTGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTGCTTCATCTACAGGTGAAGTTCA 373

RESULT 15
US-10-315-920-5
; Sequence 5, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-5
Query Match 23.5%; Score 41.6; DB 16; Length 678;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACCTACTGTCGCGACCCC 61
Db 206 CCAAGGTGTACGTGAAGCACCACCCCGCGACATCCCGGACTACAAGAAGCTGTCTTCCCG 265
Qy 62 TGGCGGACAAACCCGGAGCTGCATCAAGATCTCCAGAGACGCGGTGCACTACGGCTACT 121
Db 266 AGGCTTCAAGTGGAGCGCGTGTATGAACCTTCGAGGACGGCGGTGGCGACCGTGACCC 325
Qy 122 GCTACGGCTTCCAGTGTGGTGGAGTTCCTTGAAGGACGAGAAGCTCA 169
Db 326 AGGACTCTCCTCCGACGAGCGCTGCTTCATCTACAGGTGAAGTTCA 373

RESULT 16
US-10-006-922-35
; Sequence 35, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
```

```
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; FILE REFERENCE: Methods for Using the Same
; CURRENT APPLICATION NUMBER: US/10/006,922
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-35
```

```
Query Match 23.5%; Score 41.6; DB 13; Length 681;
Best Local Similarity 53.0%; Pred. No. 0.003;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 2 CCGACGTGCCAGGAACTACCCGCTGGACAGCTCGGACAAACACCTACCTGTGGGCACCCC 61
Db 209 CCAAGGTGTACGTGAAGACACCCCGCGACATCCCGACTACAAGAAGCTGTCTTCCCG 268

QY 62 TGGGCGACAAACCCGACTGCATCAAGATCTGCCAGAACGACGCGCTGACTAGGCTACT 121
Db 269 AGGGCTTCAAGTGGGAGCGGTGATGAACCTTCGAGGACGCGCGTGGCGACCGTGACCC 328

QY 122 GCTACCGGTTCCAGTGTGGTGGAGTTCCTGGAAGGACGAGAACGTCA 169
Db 329 AAGACTCTCCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCA 376
```

Search completed: July 30, 2005, 20:32:31  
Job time : 405.323 secs

```
RESULT 15
US-10-006-922-37
; Sequence 37, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; FILE REFERENCE: Methods for Using the Same
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 681
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:50:29 ; Search time 1939.04 Seconds  
(without alignments)  
3474.585 Million cell updates/sec

Title: US-10-617-978-17\_COPY\_73\_249  
Perfect score: 177  
Sequence: 1 gccagctgccagggaacta.....acgagaactcaagggtgtga 177

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_hic.\*
- 4: gb\_est3.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_gss1.\*
- 9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	24.9	786	7	CN147592	WOUND1_50
2	43.2	24.4	632	2	BE262829	BE262829 601151586
3	42.6	24.1	466	7	CF846529	CF846529 psHB038xO
4	42.6	24.1	684	7	CF847742	CF847742 DSMA002xG
5	42	23.7	778	4	EG836816	EG836816 Zm08_04h1
6	40.4	22.8	446	5	BQ294408	BQ294408 1091028H0
7	40.4	22.8	513	5	BQ779473	BQ779473 946120G01
8	40.4	22.8	571	4	BI478941	BI478941 949071B07
9	40.4	22.8	617	6	CA828471	CA828471 1114028F0
10	40.4	22.8	620	6	CB886436	CB886436 3529_1_95
11	40.4	22.8	624	5	BU098969	BU098969 946160D07
12	40.4	22.8	624	5	BU098975	BU098975 946160E01
13	40.4	22.8	652	5	BU098330	BU098330 946134C03
14	40.4	22.8	666	5	BU499682	BU499682 946178C07
15	40.4	22.8	668	6	CD001260	CD001260 3529_1_95
16	40.4	22.8	694	7	CF635183	CF635183 zmrw00_0
17	40.4	22.8	702	6	CD446144	CD446144 EL01T0208
18	40.4	22.8	780	7	CK371039	CK371039 zmrw005
19	40.4	22.8	834	6	CD437439	CD437439 EL01N0501
20	40.4	22.8	835	6	CD436507	CD436507 EL01N0358
21	40.4	22.8	886	6	CD435281	CD435281 EL01N0358
22	40.4	22.8	1122	3	AY103942	AY103942 Zea mays
23	40.2	22.7	1080	6	CD504256	CD504256 CDA66-H11
24	39.8	22.5	201	2	BE294986	BE294986 601175313

25	39.8	22.5	230	4	BF982489	BF982489 602305172
26	39.8	22.5	241	2	BF305332	BF305332 601892806
27	39.8	22.5	292	4	BF981884	BF981884 602309883
28	39.8	22.5	395	7	CF144021	CF144021 UI-HF-BPO
29	39.8	22.5	463	2	AW500841	AW500841 UI-HF-BPO
30	39.8	22.5	474	7	CN278122	CN278122 328743034
31	39.8	22.5	583	5	BP321118	BP321118 BP321118
32	39.8	22.5	589	5	BP266195	BP266195 BP266195
33	39.8	22.5	593	6	CD611367	CD611367 56087962J
34	39.8	22.5	613	4	BI261161	BI261161 602972621
35	39.8	22.5	697	1	AA522689	AA522689 ni39c06.8
36	39.8	22.5	709	7	CN278116	CN278116 170004708
37	39.8	22.5	741	9	CG259892	CG259892 QGVZ77TH
38	39.8	22.5	745	1	AUI32340	AUI32340 AUI32340
39	39.8	22.5	745	1	AUI39517	AUI39517 AUI39517
40	39.8	22.5	764	4	BG575900	BG575900 602598327
41	39.8	22.5	766	1	AUI30388	AUI30388 AUI30388
42	39.8	22.5	830	4	BI333060	BI333060 602996324
43	39.8	22.5	845	4	BM456185	BM456185 AGENCOURT
44	39.8	22.5	874	5	EU931387	EU931387 AGENCOURT
45	39.8	22.5	942	5	EX405398	EX405398 EX405398

ALIGNMENTS

RESULT 1  
CN147592  
LOCUS  
DEFINITION WOUND1\_50\_E10\_g1\_A002 Wounded leaves Sorghum bicolor cDNA clone  
786 bp mRNA linear EST 01-APR-2004  
WOUND1\_50\_E10\_A002 5', mRNA sequence.  
ACCESSION CN147592  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Sorghum bicolor (sorghum)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Sorghum.  
1 (bases 1 to 786)  
Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,  
Sun, F., Sullivan, R., Brady, J., Eastman, A., Miller, V., Gonzalez, M.,  
Anfuso, C., Chhabra, D., Johnson, H., Kamran, D. and Pratt, L.H.  
A Sorghum EST database: mechanically damaged and methyl  
jasmonate-treated leaves  
Unpublished (2003)  
Other ESTs: WOUND1\_50\_E10\_b1\_A002  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
the Human Genome Center, University of Tokyo Institute of Medical  
Science; plant material and RNA prepared at Texas A & M University;  
sequencing done in the Laboratory for Genomics and Bioinformatics,  
University of Georgia. Sequence ends have been trimmed to exclude  
vector and regions below Phred quality 16. Three-prime sequences  
are presented as their reverse complement and have been trimmed to  
exclude polyA.  
Seq primer: Sugs (CTTCTGCTCTAAAGCTCGG).  
Location/Qualifiers  
1..786  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="WOUND1\_50\_E10\_A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Wounded leaves"  
/note="Organ: Leaf; Vector: pME185-FLJ3; Site\_1: XhoI;

FEATURES  
source  
1..786  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="WOUND1\_50\_E10\_A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Wounded leaves"  
/note="Organ: Leaf; Vector: pME185-FLJ3; Site\_1: XhoI;

Site\_2: XhoI; The library was prepared from polyA+ RNA harvested from 8-day-old hydroponically grown, BTx623 sorghum seedlings. For some plants, one-half of the second leaf was crushed without damaging the midvein. For others, methyl jasmonate was added to the growth medium to a final concentration of 100  $\mu$ M. Leaves were harvested 3 and 27 hr after treatment and pooled. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-F13 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

## ORIGIN

Query Match 24.9%; Score 44; DB 7; Length 786;  
 Best Local Similarity 57.1%; Pred. No. 0.12;  
 Matches 80; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 30 CAGCTGGACAAACCTACTGTGGCGACCCCTGGGGGACAAACCGGAGTGCATCAAGAT 89  
 Db 350 CGGCGGGGACAACTCAACCCCATCTGGCCCATCGTCAACGACACACGCTCAAGTCTCAGGAC 409

Qy 90 CTGCCAGAACGCGGTGCTACGCTACTGCTACGCGTTCAGTGTTCAGTGGCGAGTT 149  
 Db 410 CTTCTACTCGACTTGGCGAGGACAACTACTTCTGTCGGGACAGCGTGGTTCGAGTC 469

Qy 150 CCTGAAGGACGAGAACTCA 169  
 Db 470 CATCATCCACGCCATCGCA 489

## RESULT 2

BE262829  
 LOCUS 601151586F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3507620 5',  
 DEFINITION mRNA sequence.

ACCESSION BE262829.1 GI:9136188

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L1CM187 row: 1 column: 21

High quality sequence start: 27

High quality sequence stop: 616.

Location/Qualifiers

## FEATURES

source

1..632

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3507620"

/tissue\_type="neuroblastoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_19"

/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site\_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 24.4%; Score 43.2; DB 2; Length 632;  
 Best Local Similarity 54.4%; Pred. No. 0.19; 73; Indels 0; Gaps 0;  
 Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1 GCCGAGCTGCCAGGGAACTACCCGCTGGGACAGCTCGGACAAACACCTACCTGTGGCACCC 60  
 Db 184 GGCGGTGGGCTGAGCGCAAGAGCAAGCCCAACATCACTACGACAAGCTGAGCGCGGCC 243

Qy 61 CTGGGCGACAAACCCGAGCTGCTCAAGATTCGCCAAGACACCGGCTCGACTACGGGTAC 120  
 Db 244 CTGCGCTACTACTACGACAAAGAACATCATGAGCAAGGTGCATGGCAAGCGCTACGCGCTAC 303

Qy 121 TGTCTACGCTTCCAGTGTGGTGGCGAGTTCCTGAAGGAGC 160  
 Db 304 CGCTTCGACTTTCAGGGCCTGGCGCAGTGCCTGCCAGCCG 343

## RESULT 3

CF846529

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CF846529.1 GI:38062183

EST.

Phytophthora sojae

Phytophthora sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phytophthora.

1 (bases 1 to 466)

Tyler, B.

Not Published

Unpublished (2003)

Contact: Tyler B

Tyler lab

VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmtylev@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 038 row: O column: 20

Seq primer: BK reverse primer

High quality sequence stop: 466.

Location/Qualifiers

1..466

/organism="Phytophthora sojae"

/mol\_type="mRNA"

/db\_xref="taxon:67593"

/clone="GH038020"

/tissue\_type="mycelium"

/cell\_line="P6497"

/dev\_stage="48 hr. post infection stage"

/lab\_host="Soybean plant"

/clone\_lib="USDA-IFAFS:Expression of Phytophthora sojae

genes during infection and propagation\_sHB"

/notes="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 24.1%; Score 42.6; DB 7; Length 466;  
 Best Local Similarity 59.5%; Pred. No. 0.27;  
 Matches 72; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 43 ACCTACTGTGGCCACCCCTGGGCGCAACACCGGAGTGCATCAAGATCTCCAGAAGCAC 102  
 Db 260 ATCCAGCGCACCCCGGCTGTGCAAGAGGGCGGCTTCAGCAAGTTCATCGGCAGAAC 319

Qy 103 GGCGTCGACTACGGCTACTGCTACGGTTCACGTTCTTGGTGGCGAGTTCTCTGAAGGACGAG 162

Db 320 CCAGCAGCTACGACGGCGGTAGCTGCTGCGGCTACGAGGTCCCGAAGAGCGAC 379  
 Qy 163 A 163  
 Db 380 A 380

RESULT 4  
 CF847742  
 LOCUS  
 DEFINITION pMA002x02f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sMA Phytophthora sojae cDNA clone sMA002G02 5, mRNA sequence.

ACCESSION CF847742  
 VERSION CF847742.1 GI:38063396  
 KEYWORDS EST.  
 SOURCE Phytophthora sojae  
 ORGANISM Phytophthora sojae  
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.

REFERENCE 1 (bases 1 to 684)  
 Tyler, B.  
 Tyler, B. Not Published  
 TITLE Unpublished (2003)  
 JOURNAL Contact: Tyler B  
 COMMENT Tyler lab

FEATURES  
 source  
 1. 684  
 /organism="Phytophthora sojae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:67593"  
 /clone="sMA002G02"  
 /tissue\_type="mycelium"  
 /cell\_line="P6497"  
 /dev\_stage="mycelium"  
 /lab\_host="synthetic medium"  
 /clone\_lib="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sMA"  
 /note="vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
 Query Match 24.1%; Score 42.6; DB 7; Length 684;  
 Best Local Similarity 59.5%; Pred. No. 0.28; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

Qy 43 ACCTACCTGTGCGCACCCCTGGGCGACACCCGAGCTGCATCAAGATCTGCCAAGCAC 102  
 Db 518 ATCCACGCGCACCCGCGGCTGTGCAAGAGCGCGCCCTTCAGCAAGTTTCATCGGCAAGAC 577

Qy 103 GCGGTCGACTACGCTACTGCTAGCGGTTTCAGTGTGGTGGCGAGTTCTCGAAGGACGAG 162  
 Db 578 CCCAGCAGCTACGACGCGGCTAGCTGAGCGCTGTGCGGCTACGAGGTCCCGAAGAGCGAC 637

Qy 163 A 163  
 Db 638 A 638

RESULT 5  
 BG836816  
 LOCUS  
 DEFINITION Zm08\_04h11\_A

Db 320 CCAGCAGCTACGACGGCGGTAGCTGCTGCGGCTACGAGGTCCCGAAGAGCGAC 379  
 Qy 163 A 163  
 Db 380 A 380

RESULT 4  
 CF847742  
 LOCUS  
 DEFINITION pMA002x02f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sMA Phytophthora sojae cDNA clone sMA002G02 5, mRNA sequence.

ACCESSION CF847742  
 VERSION CF847742.1 GI:38063396  
 KEYWORDS EST.  
 SOURCE Phytophthora sojae  
 ORGANISM Phytophthora sojae  
 Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.

REFERENCE 1 (bases 1 to 684)  
 Tyler, B.  
 Tyler, B. Not Published  
 TITLE Unpublished (2003)  
 JOURNAL Contact: Tyler B  
 COMMENT Tyler lab

FEATURES  
 source  
 1. 684  
 /organism="Phytophthora sojae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:67593"  
 /clone="sMA002G02"  
 /tissue\_type="mycelium"  
 /cell\_line="P6497"  
 /dev\_stage="mycelium"  
 /lab\_host="synthetic medium"  
 /clone\_lib="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_sMA"  
 /note="vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
 Query Match 24.1%; Score 42.6; DB 7; Length 684;  
 Best Local Similarity 59.5%; Pred. No. 0.28; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0

Qy 43 ACCTACCTGTGCGCACCCCTGGGCGACACCCGAGCTGCATCAAGATCTGCCAAGCAC 102  
 Db 518 ATCCACGCGCACCCGCGGCTGTGCAAGAGCGCGCCCTTCAGCAAGTTTCATCGGCAAGAC 577

Qy 103 GCGGTCGACTACGCTACTGCTAGCGGTTTCAGTGTGGTGGCGAGTTCTCGAAGGACGAG 162  
 Db 578 CCCAGCAGCTACGACGCGGCTAGCTGAGCGCTGTGCGGCTACGAGGTCCCGAAGAGCGAC 637

Qy 163 A 163  
 Db 638 A 638

RESULT 5  
 BG836816  
 LOCUS  
 DEFINITION Zm08\_04h11\_A

Zm08\_AAPC\_ECORC\_Fusarium\_graminearum\_inoculated\_corn\_ear Zea mays  
 cDNA\_clone Zm08\_04h11, mRNA sequence.

ACCESSION BG836816  
 VERSION BG836816.1 GI:14203139  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 778)  
 Harris, L.J., Balcerzak, M., Allard, S., Saparno, A., Couroux, P., De Moors, A., Hattori, J.I., Ouellet, T., Robert, L.S., Singh, J.A., Sprott, D., and Linker, N.A.  
 TITLE Expressed Sequence Tags from Developing Maize Kernels Six Days after Silk Channel Inoculation with Fusarium graminearum  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Harris, Linda J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Bldg. 21, Central Experimental Farm, Ottawa, Ontario, KIA 0C6, CANADA  
 Tel: (613) 759-1314  
 Fax: (613) 759-6566  
 Email: harrielj@agr.gc.ca

FEATURES  
 source  
 1. 778  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="CO430"  
 /db\_xref="taxon:4577"  
 /clone="Zm08\_04h11"  
 /tissue\_type="Developing kernels (silb crossed)"  
 /dev\_stage="10-11 days post-silk emergence"  
 /clone\_lib="Zm08\_AAPC\_ECORC\_Fusarium\_graminearum\_inoculated\_corn\_ear"  
 /note="vector: Bluescript SK+/XhoI-EcoRI; Site\_1: EcoRI; Site\_2: XhoI; Field-grown maize ears were silk channel-inoculated in the morning (~10 am) with 1 ml of a Fusarium graminearum macroconidial suspension (500,000 spores/ml) and whole ears were collected and immediately frozen in liquid nitrogen 6 days later."

ORIGIN  
 Query Match 23.7%; Score 42; DB 4; Length 778;  
 Best Local Similarity 55.5%; Pred. No. 0.41; Mismatches 81; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 29 ACAGCTCGGACACACCTACCTGTGCGCACCCCTGGGCGACACCCGAGCTGCATCAAGA 88  
 Db 128 ACAAGGTGTACAACGGCAGCTTGGACAAAGCGGCTGCGGCGTACCCGCGGCACTCTGCGTGC 187

Qy 89 TCTGCCAGAGCAGCGGCTCGACTACGCTACTCTAGCGCTTCAGGTTTGTGGTGGCAGT 148  
 Db 188 TCATCCAGAGAGTCCCGACCGGACCGGACCGGACCGCTAGAGGCCATCTACAGCTTCTACT 247

Qy 149 TCCTGAAGGACGAGAACGTCAAGGTG 174  
 Db 248 TCGGAGACTACGGCCACATCTCGGTG 273

RESULT 6  
 BQ294408  
 LOCUS  
 DEFINITION 1091028H08.y2 1091 - Immature ear with common ESTs screened by Schmidt lab Zea mays cDNA, mRNA sequence.

ACCESSION BQ294408  
 VERSION BQ294408.1 GI:20803358  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD



	ORIGIN	selected."
Query Match	22.8%; Score 40.4; DB 6; Length 617;	
Best Local Similarity	54.8%; Pred. No. 1.1;	
Matches 80; Conservative	0; Mismatches 66; Indels 0; Gaps 0;	
Qy	29 ACAGCTCGGACACACTTCTGTGGCACCCCTGGCGACAAACCCGGACTGCATCAAGA	88
Db	327 ACAAGTGTTACAACGGCAGCCTGGACAAGCGGTGGGGCTACCGGGGGCATCTCGGTGC	386
Qy	89 TCTGCCAAGAAGCAGCGGCTCGACTACGGCTACTGTCTACGGTTCAGTGTGTGGTCGAGT	148
Db	387 TCATCCAGCAGCTGCCCGACCAGCGGGACCGCTACGAGGCCATCTCAGCTTCTACT	446

QY	149	TCTGTGAAGGACGAGAAGTCAAGTG	174
DB	447	TCGGAGACTACGCCCATCTCGGTG	472

  

RESULT 10
CB886436/c

LOCUS	CB886436	620 bp	mRNA	linear	EST 23-APR-2003
DEFINITION	3529..1..95.._F01.x.1 3529 - 2 mm ear tissue from Schmidt and Hake				
	labs Zea mays cDNA, mRNA sequence.				
ACCESSION	CB886436				
VERSION	CB886436.1	GI:30088231			
KEYWORDS	EST.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota;	Viridiplantae;	Streptophyta;	Embryophyta;	Tracheophytes;
	Spermatophyta;	Magnoliophyta;	Liliopsida;	Poales;	Poaceae;
	PACCAD				

REFERENCE	1 (pages 1 to 620)
AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University
COMMENT	Unpublished (1999) Contact: walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 3529_1_95_1 row: F column: 01.
FEATURES	Location/Qualifiers
source	1. .620

```

/organism="sea_maya"
/mol_type="mrna"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="ear"
/dev_stage="2 mm"
/lab_host="E. coli XL0LR"
/clone_lib="3529 - 2 mm ear tissue from Schmidt and Hake
labs"
/note="Organ: ear; Vector: PAD-GAL4-2.1; Site.1: EcoRI;
Site.2: XhoI; RNA isolated by Hake lab. 1 million pfu
amplified. Ampicillin is the selection marker."

```

[illegible]





```
JOURNAL COMMENT
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946134 row: C column: 03.

FEATURES
source
1..652
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XL0LR"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/notes="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

ORIGIN
Query Match 22.8%; Score 40.4; DB 5; Length 652;
Best Local Similarity 54.8%; Pred. No. 1.1;
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 29 ACAGCTCGGACACACCTACCTGTGCGCACCCCTGGGCGACAAACCGGACTGTCATCAAGA 88
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 ACAGGTGTACACGGCAGCTGTGACAGCGGTGGGCGTACCGCGGGCATCTCGGTGC 419
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 89 TCTGCCAGACGCGGCTGACTACGGCTACTCTACGGCTTCTCAGTTCAGTGTGGTGCAGT 148
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 TCATCCAGCAGCTCCCGACCGCAACGCGACCGCTACGAGGCCATCTACAGCTTCTACT 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 149 TCCTGAGGACGAGACGTCAGGTG 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 TCGGAGACTACGGCCACATCTCGGTG 505
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
BU499682
LOCUS 946178C07.y1 946 - tassels primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BU499682
VERSION BU499682.1 GI:22819592
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 666)

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946178 row: C column: 07.

FEATURES
source
1..666
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XL0LR"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/notes="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

ORIGIN
Query Match 22.8%; Score 40.4; DB 5; Length 666;
Best Local Similarity 54.8%; Pred. No. 1.1;
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 29 ACAGCTCGGACACACCTACCTGTGCGCACCCCTGGGCGACAAACCGGACTGTCATCAAGA 88
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 ACAGGTGTACACGGCAGCTGTGACAGCGGTGGGCGTACCGCGGGCATCTCGGTGC 390
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 89 TCTGCCAGACGCGGCTGACTACGGCTACTCTACGGCTTCTCAGTTCAGTGTGGTGCAGT 148
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 TCATCCAGCAGCTCCCGACCGCAACGCGACCGCTACGAGGCCATCTACAGCTTCTACT 450
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 149 TCCTGAGGACGAGACGTCAGGTG 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 TCGGAGACTACGGCCACATCTCGGTG 476
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
CD001260
LOCUS 3529.1_95.1_F01.y.1 3529 - 2 mm ear tissue from Schmidt and Hake
DEFINITION labs Zea mays cDNA, mRNA sequence.
ACCESSION CD001260
VERSION CD001260.1 GI:30306587
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 668)

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 3529.1_95.1 row: F column: 01.

FEATURES
source
1..668
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="ear"
/dev_stage="2 mm"
/lab_host="E. coli XL0LR"
/clone_lib="3529 - 2 mm ear tissue from Schmidt and Hake
labs"
/notes="Organ: ear; Vector: PAD-GAL4-2.1; Site 1: EcoRI;
```

Site 2: XhoI; RNA isolated by Hake lab. 1 million pfu amplified. Ampicillin is the selection marker."

ORIGIN

Query Match	22.8%	Score 40.4;	DB 6;	Length 668;
Best Local Similarity	54.8%	Pred. No. 1.1;	Indels 0;	Gaps 0;
Matches	80;	Conservative 0;	Mismatches 66;	

  

Qy	29	ACAGCTCGGACAAACACCTTACCTGTGCGCACCCCTGGGCGGACAAACCGGACTGCATCAAGA	88
Db	373	ACAAGTGTACACGGCAGCTGGACAAGCGGCTGGGCGTCACCGGGGATCTGGGTGC	432
Qy	89	TCTGCCAGAACGACGGGCTCGACTACGGCTACTGCTACGGCTTCCAGTGTGGTGGAGT	148
Db	433	TCATCCAGCACGTCCCGGACCGCAACCGGCGACCGCTACGAGGCGATCTACAGCTTCTACT	492
Qy	149	TCCTGAGGACGAGACGTCAAGTG	174
Db	493	TCGAGACTACGGCCACATCTCGGTG	518

Search completed: July 30, 2005, 20:16:27  
Job time : 1947.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2005, 15:39:09 ; Search time 1277.13 Seconds  
(without alignments)  
6791.361 Million cell updates/sec

Title: US-10-617-978-14\_COPY\_62\_240  
Perfect score: 179  
Sequence: 1 cgctgacgtcccggaac.....atgagaacgtgaaggtctga 179

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hug.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.2	32.0	354	3 AF151798	AF151798 Mesobuthus
2	57.2	32.0	469	3 AY282464	AY282464 Mesobuthu
3	42.2	23.6	270	6 AR566415	AR566415 Sequence
4	42.2	23.6	270	6 AX061772	AX061772 Sequence
5	40.8	22.8	270	6 AR566414	AR566414 Sequence
6	40.8	22.8	270	6 AR566416	AR566416 Sequence
7	40.8	22.8	270	6 AX061770	AX061770 Sequence
8	40.8	22.8	270	6 AX061774	AX061774 Sequence
9	38	21.2	270	6 AR566413	AR566413 Sequence
10	38	21.2	270	6 AX061768	AX061768 Sequence
11	38	21.2	316	3 AF491133	AF491133 Centruroi
12	36.8	20.6	192	3 AY351304	AY351304 Centruroi
13	36.8	20.6	192	3 AY351305	AY351305 Centruroi
14	36.8	20.6	192	3 AY351306	AY351306 Centruroi
15	36.8	20.6	192	3 AY351307	AY351307 Centruroi
16	36.8	20.6	319	3 AF338459	AF338459 Centruroi
17	36.8	20.6	320	3 AF338461	AF338461 Centruroi
18	36.8	20.6	323	3 AF338454	AF338454 Centruroi
19	36.8	20.6	323	3 AF338455	AF338455 Centruroi

20	36.8	20.6	323	3 AF338456	AF338456 Centruroi
21	36.8	20.6	323	3 AF338457	AF338457 Centruroi
22	36.8	20.6	323	3 AF338458	AF338458 Centruroi
23	36.8	20.6	323	3 AF338462	AF338462 Centruroi
24	36.8	20.6	323	3 AF338463	AF338463 Centruroi
25	36.8	20.6	323	3 AF338464	AF338464 Centruroi
26	36.4	20.3	273	3 AF338471	AF338471 Centruroi
27	35.2	19.7	192	3 AY351301	AY351301 Centruroi
28	35.2	19.7	323	3 AF338453	AF338453 Centruroi
c	35	19.6	110000	2 EX649244	Continuation (4 of
c	35	19.6	225818	10 AL833792	AL833792 Mouse DNA
31	34.8	19.4	252	3 AY649869	AY649869 Centruroi
32	34.8	19.4	254	3 AY649866	AY649866 Centruroi
33	34.8	19.4	2993	5 AF032382	AF032382 Xenopus 1
34	34.6	19.3	261	3 AY649863	AY649863 Centruroi
35	34.6	19.3	261	3 AY649865	AY649865 Centruroi
36	34.6	19.3	314	3 AY649860	AY649860 Centruroi
37	34.6	19.3	315	3 AY649859	AY649859 Centruroi
c	34.4	19.2	191952	2 AC112424	AC112424 Rattus no
39	34.4	19.2	215115	2 AC094688	AC094688 Rattus no
40	33.8	18.9	265229	2 AC121468	AC121468 Rattus no
41	33.6	18.8	261	3 AY649861	AY649861 Centruroi
42	33.6	18.8	261	3 AY649864	AY649864 Centruroi
43	33.6	18.8	314	3 AY649862	AY649862 Centruroi
44	33.4	18.7	314	3 AF338450	AF338450 Centruroi
45	33	18.4	188640	2 AC146153	AC146153 Pan trogl

#### ALIGNMENTS

RESULT 1	AF151798	Mesobuthus martensii	354 bp	mRNA	linear	INV 06-OCT-2004
LOCUS	AF151798	cds.				
DEFINITION	AF151798	Mesobuthus martensii	insect beta-neurotoxin (bt)	mRNA, complete		
ACCESSION	AF151798					
VERSION	AF151798.2	GI:53828929				
KEYWORDS		Mesobuthus martensii (Buthus martensii)				
SOURCE		Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthida; Buthoidea; Buthidae; Mesobuthus.				
REFERENCE		1 (bases 1 to 354)				
AUTHORS		Zeng,X.-C., Li,W.-X. and Zu,S.-Y.				
TITLE		A novel cDNA sequence encoding the precursor of a new type of insect beta-neurotoxin, BmKBT from Chinese scorpion Buthus martensii Karsch				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 354)				
AUTHORS		Li,W.-X., Zeng,X.-C. and Zu,S.-Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (15-MAY-1999) Department of Virology and Molecular Biology, Virology Institution of Wuhan University, Luojia Street, Wuhan, Hubei 430072, P.R. China				
REFERENCE		3 (bases 1 to 354)				
AUTHORS		Li,W.-X., Zeng,X.-C. and Zu,S.-Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (05-OCT-2004) Department of Virology and Molecular Biology, Virology Institution of Wuhan University, Luojia Street, Wuhan, Hubei 430072, P.R. China				
REMARK		Sequence update by submitter				
COMMENT		On Oct 6, 2004 this sequence version replaced gi:30844240.				
FEATURES		Location/Qualifiers				
source		1..354				
		/organism="Mesobuthus martensii"				
		/mol_type="mRNA"				
		/db_xref="taxon:34649"				
		/tissue_type="venom gland"				
		/note="Authority: Buthus martensii Karsch"				
gene		1..354				
CDS		/gene="bt"				
		12..254				

```
/gene="bt"  
/site="BmkBT2"  
/codon_start=1  
/product=insect beta-neurotoxin"  
/protein_id="AAPA1418.2"  
/db_xref="GI:53828930"  
/translation="MKFKLLPGMEVILFSLMGSTIRGDDDFGNVPTNAYGNKYYCTILG  
ENYCKLKLHGVTYGVCYSNCWCEKLDKDTI"
```

Query Match	32.0%;	Score 57.2;	DB 3;	Length 354;
Best Local Similarity	58.0%;	Pred. No. 1.2e-07;		
Matches 101;	Conservative	0;	Mismatches 73;	Indels 0;
				Gaps 0;

Qy	6	GAGTCCGGGAACTACCCACITGATTCTTCGACAAATACCTACCTGTGGGCCCTTG	65
Db	81	GACATCCAGAAATATTCGACAAATGCTTACCGTAACAAATATATTTGTACGATTTG	140
Qy	66	GGAGATAATCCGGACTCGATTAGACTGTGCAGAAAACCGTGTGGATTACGGGTATTGC	125
Db	141	GGAGAAATGAATATTTAGAGAAATATGTAATTTGCATGGATTACTTATGTTATTGT	200
Qy	126	TACGCCTTCCAATGCTGGTGTGAATTTCTGAAGGATCAGACCGTGAAGGTCTGA	179
Db	201	TACAAATTCGAGATGTTGGTGCAGAAAATTTGCAAGATAAAGACGTGCACCAATTGA	254

[illegible]

```

/ country= "China"
/ note= "authority: Mesobuthus martensii Karsch"
CDS
38..280
/codon_start=1
/product= "toxin KBT precursor"
/protein_id= "AAQ22733.1"
/db_xref= "GI:33590393"
/translation= "MMKFLVFGMIVILFSLMGIRGDDPGNYPTNAYGNKYKTYTILG
ENEYCRKICKLHGVTYGYCYNSRCWCKELEDKQVTI"
38..103
sig_peptide
104..277
mat_peptide
/product= "toxin KBT"
329..334
polyA_signal
ORIGIN

```

Query Match	32.0%	Score 57.2	DB 3	Length 469
Best Local Similarity	58.0%	Pred. No. 1.2e-07		
Matches 101	Conservative	0	Mismatches 73	Indels 0
				Gaps 0

[illegible][illegible]

```

UNCLASSIFIED:
1 (bases 1 to 270)
Herrmann, R. and Wong, J.F.
Scorpion toxins
Patent: US 6768002-A 34 27-Jul-2004;
Location/Qualifiers
1. .270
/organism="unknown"
/mol_type="genomic DNA"

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

```

Query Match	23.6%	Score 42.2;	DB 6;	Length 270;
Best Local Similarity	61.3%	Pred. No. 0.011;		
Matches 68;	Conservative	0;	Mismatches 43;	Indels 0;
				Gaps 0;

[illegible][illegible]

```

ORIGIN
/seq_xref1="XaXon:6863"

Query Match      23.6%; Score 42.2; DB 6; Length 270;
Best Local Similarity 61.3%; Pred. No. 0.011;
Matches 68; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

```
Qy 67 GAGATAATCCGGACTGCATTAAAGATCTGCAGAACACACGGTGTGGATTACGGGTATTGCT 126
Db 128 GTATCATGATTATTGTCGGACATTTCTAAAGTACATGGAGTGAATATGCGGTATTGTT 187

Qy 127 ACGCCCTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177
Db 188 GGGTCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238

RESULT 5
LOCUS AR566414 270 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 32 from patent US 6768002.
ACCESSION AR566414
VERSION AR566414.1 GI:53983404
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 270)
TITLE Herrmann, R. and Wong, J.F.
JOURNAL Scorpion toxins
PATENT: US 6768002-A 32 27-JUL-2004;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
LOCATION/Qualifiers
source 1..270
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 22.8%; Score 40.8; DB 6; Length 270;
Best Local Similarity 61.1%; Pred. No. 0.033;
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATAATCCGGACTGCATTAAAGATCTGTCAGAACACGGTGTGGATTACGGGTATTGCTACG 129
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGGAGTGAATATGCGGTATTGTTGGG 190

Qy 130 CTTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177
Db 191 TCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238

RESULT 6
LOCUS AR566416 270 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 36 from patent US 6768002.
ACCESSION AR566416
VERSION AR566416.1 GI:53983406
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 270)
TITLE Herrmann, R. and Wong, J.F.
JOURNAL Scorpion toxins
PATENT: US 6768002-A 36 27-JUL-2004;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
LOCATION/Qualifiers
source 1..270
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 22.8%; Score 40.8; DB 6; Length 270;
Best Local Similarity 61.1%; Pred. No. 0.033;
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATAATCCGGACTGCATTAAAGTCTGTCAGAACACGGTGTGGATTACGGGTATTGCTACG 129
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGGAGTGAATATGCGGTATTGTTGGG 190

Qy 130 CTTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177
Db 191 TCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238
```

```
RESULT 7
LOCUS AX061770 270 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 32 from Patent WO0078957.
ACCESSION AX061770
VERSION AX061770.1 GI:12539857
KEYWORDS
SOURCE Hottentotta judaica
ORGANISM Hottentotta judaica
REFERENCE
AUTHORS 1
TITLE Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
JOURNAL Butthida; Butthoidea; Butthidae; Hottentotta.
PATENT: Herrmann, R., Lee, J.M. and Wong, J.F.
Scorpion toxins from buthotus judaicus
WO 0078957-A 32 28-DEC-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
LOCATION/Qualifiers
source 1..270
/organism="Hottentotta judaica"
/mol_type="unassigned DNA"
/db_xref="taxon:6863"
ORIGIN
Query Match 22.8%; Score 40.8; DB 6; Length 270;
Best Local Similarity 61.1%; Pred. No. 0.033;
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATAATCCGGACTGCATTAAAGATCTGTCAGAACACGGTGTGGATTACGGGTATTGCTACG 129
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGGAGTGAATATGCGGTATTGTTGGG 190

Qy 130 CTTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177
Db 191 TCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238

RESULT 8
LOCUS AX061774 270 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 36 from Patent WO0078957.
ACCESSION AX061774
VERSION AX061774.1 GI:12539859
KEYWORDS
SOURCE Hottentotta judaica
ORGANISM Hottentotta judaica
REFERENCE
AUTHORS 1
TITLE Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
JOURNAL Butthida; Butthoidea; Butthidae; Hottentotta.
PATENT: Herrmann, R., Lee, J.M. and Wong, J.F.
Scorpion toxins from buthotus judaicus
WO 0078957-A 36 28-DEC-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
LOCATION/Qualifiers
source 1..270
/organism="Hottentotta judaica"
/mol_type="unassigned DNA"
/db_xref="taxon:6863"
ORIGIN
Query Match 22.8%; Score 40.8; DB 6; Length 270;
Best Local Similarity 61.1%; Pred. No. 0.033;
Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 70 ATAATCCGGACTGCATTAAAGTCTGTCAGAACACGGTGTGGATTACGGGTATTGCTACG 129
Db 131 ATCATGATTATTGTCGGACATTTGTAAGTACATGGAGTGAATATGCGGTATTGTTGGG 190

Qy 130 CTTTCCAATGCTGGTGTGAATTTCTGAAGGATGAGAACGTGAAGTCT 177
Db 191 TCACCTCGTGTGGTGTGAATTTTGAAGAAGAAGACATCAATATTT 238
```

RESULT 9	AR566413	270 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	AR566413				
DEFINITION	Sequence 30 from patent US 6768002.				
ACCESSION	AR566413				
VERSION	AR566413.1	GI:53983403			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 270)				
AUTHORS	Herrmann,R. and Wong,J.F.				
TITLE	Scorpion toxins				
JOURNAL	Patent: US 6768002-A	30 27-JUL-2004;			
FEATURES	Location/Qualifiers				
source	1..270				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match	21.2%;	Score 38;	DB 6;	Length 270;	
Best Local Similarity	56.0%;	Pred. No. 0.28;	70;	Indels	3; Gaps 1;
Matches	93;	Conservative 0;	Mismatches		
Qy	12	CCGGGAACCTACCCACTTGATTTCTTCGACAAATACCTACCTGTGCGCCCTTTGGGAGAT	71		
Db	76	CCAGGAAATACCCGATATCTGTTATGTGTACGCTTTATGGATGCACAGCTTTTAATCAT	135		
Qy	72	AATCCGAGCTGCATTAAGATCTGTTCAGAAACACCGTGTGGATTACGGGTATTGCTACGCC	131		
Db	136	AAT---TATTGTGTGGACATTTGTAAGTACATGGAGTAAAGTATGGTATTGTTGGGTC	192		
Qy	132	TTCCAATGCTGGTGTGAATTTCTCGAAGGTAGAACGTGAAGGCT	177		
Db	193	ACCTCGTGTGGTGTGAATTTTGAAAAAGACATCGATATTT	238		
RESULT 10	AX061768	270 bp	DNA	linear	PAT 24-JAN-2001
LOCUS	AX061768				
DEFINITION	Sequence 30 from Patent W00078957.				
ACCESSION	AX061768				
VERSION	AX061768.1	GI:12539856			
KEYWORDS	.				
SOURCE	Hottentotta judaica				
ORGANISM	Hottentotta judaica				
REFERENCE	1				
AUTHORS	Herrmann,R., Lee,J.M. and Wong,J.F.				
TITLE	Scorpion toxins from buthotes judaicus				
JOURNAL	Patent: WO 0078957-A	30 28-DEC-2000;			
FEATURES	E.I. DU PONT DE NEMOURS AND COMPANY (US)				
source	Location/Qualifiers				
	1..270				
	/organism="Hottentotta judaica"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:6863"				
ORIGIN					
Query Match	21.2%;	Score 38;	DB 6;	Length 270;	
Best Local Similarity	56.0%;	Pred. No. 0.28;	70;	Indels	3; Gaps 1;
Matches	93;	Conservative 0;	Mismatches		
Qy	12	CCGGGAACCTACCCACTTGATTTCTTCGACAAATACCTACCTGTGCGCCCTTTGGGAGAT	71		
Db	76	CCAGGAAATACCCGATATCTGTTATGTGTACGCTTTATGGATGCACAGCTTTTAATCAT	135		
Qy	72	AATCCGAGCTGCATTAAGATCTGTTCAGAAACACCGTGTGGATTACGGGTATTGCTACGCC	131		
Db	136	AAT---TATTGTGTGGACATTTGTAAGTACATGGAGTAAAGTATGGTATTGTTGGGTC	192		

```

SOURCE      Centruroides sculpturatus (bark scorpion)
ORGANISM    Centruroides sculpturatus
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
            Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES   Location/Qualifiers
            source
            1..192
                /organism="Centruroides sculpturatus"
                /mol_type="genomic DNA"
                /isolate="CsEvlb"
                /db_xref="taxon:218467"
                <1..>192
                /product="beta-toxin"
                <1..>192
                /codon_start=1
                /product="beta-toxin"
                /protein_id="AAR08039.1"
                /db_xref="GI:38017467"
                /translation="KEGYLVKSDGCKYDCFWLGNKNEHCDLECKAKNOGSGSYGYCYAFA
                CWCEGLPESTPTPLPNKSC"

mRNA
CDS

Query Match      20.6%; Score 36.8; DB 3; Length 192;
Best Local Similarity 73.4%; Pred. No. 0.69;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 156
    |||||
Db 92 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 151

Qy 157 AGGA 160
    ||
Db 152 CCGA 155

RESULT 13
AY351305
LOCUS      Centruroides sculpturatus isolate CsEvlc beta-toxin gene, partial
DEFINITION
ACCESSION  AY351305
VERSION    AY351305.1
KEYWORDS   GI:38017468
SOURCE     Centruroides sculpturatus (bark scorpion)
ORGANISM   Centruroides sculpturatus
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
            Buthida; Buthoidea; Buthidae; Centruroides.
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
            Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES   Location/Qualifiers
            source
            1..192
                /organism="Centruroides sculpturatus"
                /mol_type="genomic DNA"
                /isolate="CsEvlc"
                /db_xref="taxon:218467"
                <1..>192
                /product="beta-toxin"
                <1..>192

mRNA
CDS

```

```

SOURCE      Centruroides sculpturatus (bark scorpion)
ORGANISM    Centruroides sculpturatus
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
            Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES   Location/Qualifiers
            source
            1..192
                /organism="Centruroides sculpturatus"
                /mol_type="genomic DNA"
                /isolate="CsEvlc"
                /db_xref="taxon:218467"
                <1..>192
                /product="beta-toxin"
                <1..>192
                /codon_start=1
                /product="beta-toxin"
                /protein_id="AAR08041.1"
                /db_xref="GI:38017471"
                /translation="KEGYLVKSDGCKYDCFWLGNKNEHCDLECKAKNOGSGSYGYCYAFA
                CWCEGLPESTPTPLPNKSC"

mRNA
CDS

Query Match      20.6%; Score 36.8; DB 3; Length 192;
Best Local Similarity 73.4%; Pred. No. 0.69;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 156
    |||||
Db 92 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 151

Qy 157 AGGA 160
    ||
Db 152 CCGA 155

RESULT 14
AY351306
LOCUS      Centruroides sculpturatus isolate CsEvlc beta-toxin gene, partial
DEFINITION
ACCESSION  AY351306
VERSION    AY351306.1
KEYWORDS   GI:38017470
SOURCE     Centruroides sculpturatus (bark scorpion)
ORGANISM   Centruroides sculpturatus
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
            Buthida; Buthoidea; Buthidae; Centruroides.
REFERENCE   1 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Alignment of beta-toxin nucleotide sequences
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 192)
AUTHORS     Zhu,S.
TITLE       Direct Submission
JOURNAL     Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van
            Evenstraat, Leuven, Flanders 3000, Belgium
FEATURES   Location/Qualifiers
            source
            1..192
                /organism="Centruroides sculpturatus"
                /mol_type="genomic DNA"
                /isolate="CsEvlc"
                /db_xref="taxon:218467"
                <1..>192
                /product="beta-toxin"
                <1..>192
                /codon_start=1
                /product="beta-toxin"
                /protein_id="AAR08041.1"
                /db_xref="GI:38017471"
                /translation="KEGYLVKSDGCKYDCFWLGNKNEHCDLECKAKNOGSGSYGYCYAFA
                CWCEGLPESTPTPLPNKSC"

mRNA
CDS

Query Match      20.6%; Score 36.8; DB 3; Length 192;
Best Local Similarity 73.4%; Pred. No. 0.69;
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 156
    |||||
Db 92 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTCCCAATGCTGGTGTGAATTTCTGA 151

Qy 157 AGGA 160
    ||
Db 152 CCGA 155

RESULT 15

```

AY351307  
LOCUS AY351307 192 bp DNA linear INV 31-DEC-2003  
DEFINITION Centruroides sculpturatus isolate Csevid beta-toxin gene, partial cds.  
ACCESSION AY351307  
VERSION AY351307.1 GI:38017472  
KEYWORDS  
ORGANISM  
Centruroides sculpturatus (bark scorpion)  
Centruroides sculpturatus  
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
Buthida; Buthoidea; Buthidae; Centruroides.  
REFERENCE  
1 (bases 1 to 192)  
Zhu, S.  
TITLE Alignment of beta-toxin nucleotide sequences  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 192)  
Zhu, S.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUL-2003) Leuven University, Lab of Toxicology, E van Evenstraat, Leuven, Flanders 3000, Belgium  
FEATURES  
source  
1..192  
/organism="Centruroides sculpturatus"  
/mol\_type="genomic DNA"  
/isolate="Csevid"  
/db\_xref="taxon:218467"  
c1\_>192  
/product="beta-toxin"  
c1\_>192  
/codon\_start=1  
/product="beta-toxin"  
/protein\_id="AAR08042.1"  
/db\_xref="GI:38017473"  
/translation="KSGYLVNSTGCKYDCFWLGNKNEHCDLECKAKNQGSGSYGYCYAFA  
CWCEGLPESTPTTFLPNKSC"  
mRNA  
CDS

ORIGIN

Query Match 20.6%; Score 36.8; DB 3; Length 192;  
Best Local Similarity 73.4%; Pred. No. 0.69;  
Matches 47; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 97 AGAAACACGGTGTGGATTACGGGTATTGCTACGCGCTTCCCAATGCTGGTGTGAATTTCTGA 156  
Db 92 AGAACCAAGGAGGTAGTTACGGGTATTGCTACGCTTTCCGATGCTGGTGGCAAGGTTTGC 151  
Qy 157 AGGA 160  
Db 152 CCGA 155

Search completed: July 30, 2005, 19:11:07  
Job time : 1280.13 secs